



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : C12N 7/00, C12P 21/06 C12Q 1/66	A1	(11) International Publication Number: WO 93/16172 (43) International Publication Date: 19 August 1993 (19.08.93)
(21) International Application Number: PCT/US93/00913 (22) International Filing Date: 2 February 1993 (02.02.93) (30) Priority data: 07/833,431 7 February 1992 (07.02.92) US (71) Applicants: ALBERT EINSTEIN COLLEGE OF MEDICINE OF YESHIVA UNIVERSITY [US/US]; Yeshiva University, 1300 Morris Park Avenue, Bronx, NY 10461 (US). UNIVERSITY OF PITTSBURGH [US/US]; Cathedral of Learning, Pittsburgh, PA 15260 (US). (72) Inventors: JACOBS, William, R., Jr. ; 163 Fordham Street, City Island, NY 10464 (US). BLOOM, Barry, R. ; 61 Summit Drive, Hastings-on-Hudson, NY 10706 (US). HATFULL, Graham, F. ; 6 Forbes Terrace, Pittsburgh, PA 15217 (US).	(74) Agent: GEORGE, Kenneth, P.; Amster, Rothstein & Eb- enstein, 90 Park Avenue, New York, NY 10016 (US). (81) Designated States: AU, BB, BG, BR, CA, FI, HU, JP, KR, LK, MG, MW, NO, PL, RO, RU, SD, SE, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, SN, TD, TG). Published <i>With international search report.</i>	
(54) Title: MYCOBACTERIAL SPECIES-SPECIFIC REPORTER MYCOBACTERIOPHAGES (57) Abstract <p>This invention relates to mycobacterial species-specific reporter mycobacteriophages (reporter mycobacteriophages), methods of producing such reporter mycobacteriophages and the use of such reporter mycobacteriophages for the rapid diagnosis of mycobacterial infection and the assessment of drug susceptibilities of mycobacterial strains in clinical samples. In particular, this invention is directed to the production and use of luciferase reporter mycobacteriophages to diagnose tuberculosis. The mycobacterial species-specific reporter mycobacteriophages comprise mycobacterial species-specific mycobacteriophages which contain reporter genes and transcriptional promoters therein. When the reporter mycobacteriophages are incubated with clinical samples which may contain the mycobacteria of interest, the gene product of the reporter genes will be expressed if the sample contains the mycobacteria of interest, thereby diagnosing mycobacterial infection.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SK	Slovak Republic
CI	Côte d'Ivoire	LI	Liechtenstein	SN	Senegal
CM	Cameroon	LK	Sri Lanka	SU	Soviet Union
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	MC	Monaco	TG	Togo
DE	Germany	MG	Madagascar	UA	Ukraine
DK	Denmark	ML	Mali	US	United States of America
ES	Spain	MN	Mongolia	VN	Viet Nam
FI	Finland				

MYCOBACTERIAL SPECIES-SPECIFIC REPORTER
MYCOBACTERIOPHAGES

STATEMENT OF GOVERNMENT INTEREST

This invention was made with government support under NIH Grant Number AI26170.

FIELD OF THE INVENTION

This invention relates to mycobacterial
5 species-specific reporter mycobacteriophages (reporter
mycobacteriophages), methods of making such
reporter mycobacteriophages, and the use of such
reporter mycobacteriophages, for example, to rapidly
diagnose mycobacterial infection and to assess drug
10 susceptibilities of mycobacterial strains in clinical
samples. Specifically, this invention relates to the
use of mycobacterial species-specific luciferase
reporter mycobacteriophages to diagnose tuberculosis
and to assess the drug susceptibilities of the various
15 strains of Mycobacterium tuberculosis
(M. tuberculosis).

To produce the mycobacterial species-specific
reporter mycobacteriophages of the invention,
transcriptional promoters and reporter genes are
20 introduced into the genomes of mycobacterial

species-specific mycobacteriophages. These reporter genes may be the genes for luciferase or the β -galactosidase gene, and provide the DNA which encodes production of a gene product. The reporter

5 mycobacteriophages may be used by incubating same with samples which may contain the specific mycobacteria of interest. If the mycobacteria of interest is present, then the reporter mycobacteriophages introduce the recombinant nucleic acids which encode expression of

10 the gene product into the mycobacteria of interest, and the mycobacteria then express the gene product. The expressed reporter gene product may be detected by a suitable assay, for example, through the detection of photons or the conversion of an easily assayable

15 chemical reaction. The presence of such gene product indicates that the sample contains the mycobacteria of interest, and hence the mycobacterial species-specific reporter mycobacteriophages may be used to detect and thereby diagnose the specific mycobacterial

20 infection. In addition, since signals may not be generated by cells which are not metabolically active in the presence of antibiotics, the mycobacteria species-specific reporter mycobacteriophages of this invention may be used to assess the drug

25 susceptibilities of various strains of mycobacteria. If antibiotic drugs are added to the sample containing the reporter mycobacteriophages and the gene product

is detected, the mycobacteria is metabolically active and hence resistant to the antibiotic drug.

BACKGROUND OF THE INVENTION

In 1990, there was a 10% increase in the
5 incidence of tuberculosis in the United States. In addition, there has been an increase in the appearance of clinical isolates of tuberculosis that are resistant to antibiotics used to treat the disease. This problem is exacerbated by the length of time that
10 is currently needed both to diagnose tuberculosis, and to determine the drug susceptibilities of various strains of M. tuberculosis. As a result, patients with M. tuberculosis may remain infectious for long periods of time without being treated, or may be
15 treated with a drug to which the bacterial strain is resistant. Therefore, a need has arisen in the field for a method of diagnosis of M. tuberculosis (and other mycobacterial infections) which is rapid, sensitive and specific, which method is also capable
20 of assessing the drug susceptibilities of the various strains of M. tuberculosis and other mycobacterial strains. It is critical that a mycobacterial strain be assessed for drug resistance rapidly because a
25 patient infected with a strain of M. tuberculosis or another mycobacteria must be treated immediately with the particular antibiotic drug(s) to which the strain is not resistant, and not with antibiotic drug(s) to

which the strain is resistant, or the patient may die.

Currently, the most rapid test available for the diagnosis of M. tuberculosis is the staining of sputum samples for acid-fast bacilli, which is a tedious procedure, and which procedure has high sensitivity. Alternative methods for diagnosis require cultivation of the bacilli for approximately two to six weeks followed by classification of the cultured organism. Typical diagnostic tools include biochemical tests, analysis of mycolic acids and serotyping. All of these tests are time-consuming. More recently, the use of oligonucleotide probes and Polymerase Chain Reaction have been suggested for the identification of M. tuberculosis species. Although these methods may be useful approaches, their uses in a clinical setting have not yet been determined. Further, these methods do not distinguish between live and dead organisms, and are therefore of limited use in the determination of drug sensitivities of clinical isolates.

In addition, Mycobacterium avium (M. avium) is a mycobacteria which is often found in immunosuppressed patients. This mycobacteria is typically disseminated throughout the bodies of immunosuppressed patients, such as AIDS patients, and causes M. avium infection. Because this mycobacteria often causes death in immunosuppressed patients, it is

necessary to be able to diagnose and assess the drug susceptibilities of the various strains of M. avium.

It is therefore an object of this invention to construct broad mycobacterial host range and

5 mycobacterial species-specific reporter mycobacteriophages.

It is another object of this invention to provide mycobacterial species-specific reporter mycobacteriophages which may be used to rapidly
10 diagnose mycobacterial infections.

It is still another object of this invention to provide mycobacterial species-specific reporter mycobacteriophages which may be used to rapidly assess the drug susceptibilities of different strains of
15 mycobacteria in clinical samples.

It is yet another object of this invention to provide mycobacterial species-specific reporter mycobacteriophages wherein the reporter genes are luciferase genes, which mycobacterial species-specific
20 reporter mycobacteriophages may be used to rapidly diagnose mycobacterial infections and to rapidly assess the drug susceptibilities of various strains of mycobacteria.

It is a further object of this invention to
25 provide mycobacterial species-specific luciferase gene reporter mycobacteriophages which may be used to rapidly diagnose tuberculosis and assess the drug

susceptibilities of the various strains of M. tuberculosis.

SUMMARY OF THE INVENTION

This invention relates to broad host range and
5 mycobacterial species-specific reporter
mycobacteriophages, (reporter mycobacteriophages),
methods of producing such reporter mycobacteriophages,
and the use of such reporter mycobacteriophages to
rapidly diagnose mycobacterial infection, such as
10 M. tuberculosis, and to distinguish which strains of
the mycobacteria are drug-resistant. To produce these
reporter mycobacteriophages, reporter genes and
transcriptional promoters are introduced into the
genomes of mycobacterial species-specific
15 mycobacteriophages. The promoter and reporter
gene-containing mycobacteriophages (reporter
mycobacteriophages) are then incubated with a
clinical sample which may contain the mycobacteria of
interest, such as M. tuberculosis. The reporter
20 mycobacteriophages are specific for the mycobacteria
which is sought to be detected. The reporter
mycobacteriophages efficiently introduce the
recombinant nucleic acids which encode the expression
of the reporter gene's gene product into the
25 mycobacteria of interest, and the mycobacteria then
express the gene product. A substrate or other means

capable of allowing for the detection of the gene product is then added to the sample. If the gene product or the signal generated by the gene product is detected, the presence of the infectious mycobacteria is known, thereby diagnosing the disease. To assess drug susceptibility of mycobacteria, drugs such as antibiotics may be added to a sample containing the reporter mycobacteriophages of this invention. If the mycobacteria are susceptible to a drug after exposure to the drug, the mycobacteria will be killed.

However, drug-resistant mycobacteria will continue to be metabolically active in the presence of the drug, and will continue to express the detectable gene product of the reporter genes.

The preferred reporter genes of the present invention are the Firefly luciferase lux gene (FFlux), the luciferase lux genes of Vibrio fischeri, the luciferase lux genes of Xenorhabdus luminescens and the E. coli β -galactosidase gene (lacZ). The preferred promoters of the present invention are hsp60 and L5 gene 62 promoter, and the preferred mycobacteriophages are L5, TM4 and D56A. These reporter mycobacteriophages are preferably used for the rapid diagnosis of tuberculosis and M. avium infection, and the accurate assessment of drug susceptibilities of the various strains of M. tuberculosis and M. avium.

BRIEF DESCRIPTION OF THE DRAWINGS

The above brief description, as well as further objects and features of the present invention, will be more fully understood by reference to the

5 following detailed description of the presently preferred, albeit illustrative, embodiment of the present invention when taken in conjunction with the accompanying drawings wherein:

FIGURE 1 represents the genome organization of mycobacteriophage L5;

FIGURE 2 represents a luciferase shuttle plasmid pYUB180 wherein reporter gene FFlux is fused to the BCG hsp60 promoter;

FIGURE 3 represents the amount of luciferase activity of M. smegmatis which contains the pYUB180 shuttle plasmid and the FFlux gene;

FIGURE 4 represents the effect of various antibiotic drugs on the metabolic activity of control mycobacteria and drug resistant mycobacteria in the presence of reporter mycobacteriophages which contain luciferase reporter genes;

FIGURE 5 represents shuttle plasmid phAE39 wherein the reported gene is FFlux, the promoter is hsp60, the phage is TM4 and the cosmid is pYUB216.

FIGURE 6 represents luciferase activity of M. smegmatis cells infected with shuttle phasmids phAE39; and

FIGURE 7 represents a flow chart for cloning different promoters into TM4::lux shuttle phasmid phAE39.

DETAILED DESCRIPTION OF THE INVENTION

This invention is directed to mycobacterial species-specific reporter mycobacteriophages, (reporter mycobacteriophages), methods of producing such reporter mycobacteriophages and the use of such reporter mycobacteriophages for the rapid diagnosis of mycobacterial infections and the accurate assessment of mycobacterial drug susceptibilities.

In order to produce such reporter mycobacteriophages, mycobacterial species-specific mycobacteriophage genomes are modified by introducing therein transcriptional promoters and reporter genes whose gene product can be sensitively detected. The reporter mycobacteriophages may then be incubated with clinical samples suspected of containing the mycobacteria of interest, either directly or after culture, and the samples tested for the presence of the reporter gene product, thereby diagnosing mycobacterial infection.

The method of this invention allows for rapid diagnosis because only the amount of time necessary for the reporter mycobacteriophages to infect their host cells and the amount of time necessary for the host cells to synthesize the reporter gene product are

required to allow for diagnosis. Typically, the amount of time required for the reporter mycobacteriophages to infect their host cells and for the host cells to synthesize the reporter gene product is between ten minutes and sixteen hours.

The assessment of drug susceptibilities with the reporter mycobacteriophages of this invention is accurate because the reporter mycobacteriophages only allow for the detection of metabolically active mycobacterial organisms, the presence of which metabolic activity indicates that a drug has not killed the mycobacteria and that the mycobacteria is resistant to the drug.

To enhance diagnosis specificity, a series of similar reporter mycobacteriophages, each of which having well-defined but different specificities for mycobacterial species, is selected.

Mycobacteriophage L5, a temperate virus with a broad host-range among mycobacteria, is the most thoroughly characterized of the mycobacteriophages. L5 particles are morphologically similar to the family of phages that includes phage ϕ and contain a linear dsDNA genome with cohesive ends. The inventors have determined the DNA sequence of the entire gene as well as several gene functions. The DNA sequence of the L5 mycobacteriophage is as follows:

* * * S E Q U E N C E * * *

```

1  GGTCTGGTTAT GCGGCGGAGC CATCCTGTAC GGGTTTCCAA GTCGATCAGA GGTAGGGGGCC
61  GGCACAGAAA CCACTCACAT CAGGGCTGTG CGCCTCCAGG GCGCGTGAAC TCCCACACCC
121 CGGTGTAGTT ACATCCCGGA ATTGTCTCAG CGCCTCTCAG GCGCCTTCTC ATAAACAGTG
181 ATCTACGCCA CTCCTGACGG GTGGCTGTCA AGGATACTCA CTTTCCCTAC TAATGAGGGG
241 CTAAGAGCCC CTCTCTATAG AGCGCCGCAC AGGCGGCGCG ATAAGAGCGC CACCAGGCGC
301 TCATCTAAAG ACCGGCCTTG AAGGGCCGGT CATAGAGATC TATTGATCC GGCAACCGCC
361 GGATCTCAAG GCCGCGCCAG TCGCGGCCCG TACTCAACT TGCAATGGCA GTGCATGGCA
421 CTCGCTCGAG TGCCCACTGG AGCACTCAAC CGGGGAAGTT CGACGTTCTC AACCTGCGAA
481 TGACGTTTGA ATCGTCATCC GCGTACGAAA TCCCGATCT GCGGCGGACC GACTTCGTGC
541 CGGCCTATCT CGCGGCCTGG AATATGCCGC GTCACCGCGA TTACGCGGCC AAGAACGGCG
601 GCGCGCTGCA CTTCTTCCTT GACGATTACC GGTTTGAGAC CGCGTGGTGG TCCCCCGAGC
661 GCCTTCTEGA CCGCGTAAAG CAGGTCTGGC CTGCACTCAC GCGGATTTC AGCCTCTGGA
721 CGAACATGCC GAAGGCGGCG CAGCTATGGA ACCTCTACCG CTCCCGCTGG TGTGGCGCGT
781 ATTGGCAGTC GGAAGGAATC GAGGTGATTC CGACGGCGTG TTGGGCGACT CCCGACACGT
841 TCGATTCTTG TTTTCGACGG ATCCCGATGG GATCGACCGT CGCAATTTCT TCGATGGGCA
901 TTCGCTCTTC AAAAGTCGAC CAGGAGCTTT TCCGGTACGG ACTACGCGAA CTCATCGATC
961 GCACTCAACC GCAACTGCTT TTGGCATATG GCCAGCTTCG GCATTGCGAC ACCCTGCGAT
1021 TACCAGAGGT CCGCGAATAC CCGACCTACT GGGACAGACG ACGAAAGTGG GTAAGTCCG
1081 ATGGGAGGCC GGGGAAGTAA AGGCGGCCCG GGTCCCGGAA CCGGAGCAGC CAACCGCAGA
1141 GCGGCTGGAG CCCCCGGATC GGGCGGCTGA GCGGCGCTCG GAGGCGGGGG TGGAGCTGCA
1201 GGGAGCAGCG GAGGCGGCAA GGAACCGGCA GCGCCGGTAC CCGAGGCGTC ACCGTGGCG
1261 GCGGAAGTGG AGCCGCGGCG GTGGCGACGA GCGGCGGCTC CCGGCTGCCC CCCACCGAGC
1321 TGGAGAAGAA GCGCGGCGAA TACAACCAGA TCGCCATCGA CGCCAGAAA CAGCAGCGCG
1381 CCACCGATGA GAAGCGCGAG GCCAAGCGCA AGCAACTGAT GGATCGAGTC GAGGAGACT
1441 GGCAGGCTTT GGACCCGGAT CACCACGACG CCATCAAGGT GCGGATGGAT GACGCCATGC
1501 GGAAGATCCT CTCCGAGGAG GAGATCGTCC ACCGCACCAA GCACTTCGGC GACCTACTCG
1561 ACTCCGGTCG ACTCAAGTCG CTGTTGAGG TCGGCTTCTC AGCCGGTGGC GACACCCCGA
1621 CCGAACCGCG CCTCCTCGAG GACGCTGGT TCGGCGCAGG CAAGGTTCCC CCGATCTACT
1681 CCGCAATCGA GTTCAACGCG GCTCCGACAG CCGGCTTCGG CATGTACGGC GGCACCAAGC
1741 TCTACATGAA GGAATCGGTC AAGGACCGCG TCACCGTGAC CATCGGCGAC TCGTGATGT
1801 CGAGCTGGGA CGTATTCCCC GGCGTCTCTG GCGACGGCGT GGGGCTGTGG GCCAGCCTGT
1861 CGAAGATCGA GGGGCTGGTC GATCCGAGCA AGACCGCGCA AGAGAATATG CAGGCGGTGT
1921 ACGACTCGTT CAAGAAGTAC GGCACCTTGG ACGGCTTCAT CGAGGCGCAG ATCCACGGCG
1981 GCGTCTTGGT CGAGGACATC AAGAAGGTCT GTTTCACGCA GCGGCGGCGC CCGATCTTCA
2041 CCGATAAACT GGACGAACCT GGAATCCCGT GGGAGGTGCA GTAATGGCGC AGATGCAGGC
2101 GACACACACA ATCGAGGGGT TCCTGGCTGT CGAGGTGGCC CCTCGGGCGT TCGTCGAGA
2161 GAACGGCCAC GTACTGACCC GGCTGTCCGC CACGAAGTGG GCGGCTGGCG AGGGTCTCGA
2221 GATCCTCAAC TACGAGGGTC CAGGGACCGT CGAGGTCTCC GACGAGAAGC TCGCCGAAGC
2281 CCAGCGGGCC AGCGAGGTCT AGGCTGAACT TCGCCGCGAG GTCGCAAGG AGTGAGCTGG
2341 GCCGGCTCAG GCCGGCGACA GGAACCTACCA GAGGACTGGG AGCTGAATTA CCGGCTCCCC
2401 GTCCTTTCTG CTGCCAACTG GCTTTGCCAG ATCAACGGTC CCGGATGCGT AAGGGCCGCA
2461 ACCGATGTCT ACCACATCAA GCGCGGGAAC GACCACAGCC GGTCCAATCT GCAGGCAGCC
2521 TGCCATGTCT GTCACGGCAA GAAATCAGCC GCGGAGGGCG TAGCCCGACG GCGGGAACCT
2581 AGAGCCCGGA GGAAGCGACC ACCCGAAGCG CATCCTGGGC GTCGATAAGC GGGCCAGGTG
2641 CCGGCTCCAC CCAGGAGGTG AACAGTGGGC ACGCGAGGCC CAATCGGAAA ACGAGATGAA
2701 GAGCGGGTTC GTCGGAACAC CCCGGACAGT CCAACCGACA CGATCCAGAT GCCCGGTCTG
2761 GTGACGATCC CCGAGATGGG CGATCTAAGC CACGACCGCC GCACGCACCA GCTCGTCAAG

2821 GACATGTACG AGTCGATCAA GCAGTCGGCA GCCGTGAAGT ACTACGAGCC GACCGACTGG
2881 CAGATGGCCC GACTCGCCCT CTACACACTT AACCAGGAAC TCATCGCAGC CGAGAACAAC
2941 GGCAAGCCCG TGGGCGCGAT GAAGCTCACT GCCATCAACC AGATGCTCTC CGCGCTGCTG
3001 CTGACCGAAG GTGACCGACG CCGTACCTCGA CTCGAAGTCG AACGAGCACC CGCTGACCCG
3061 ACAGGCGGGA AGGTCTGTGA CGTGACCGAC GTGCTCAAGC AGCGCCTCGC CAAGGCGAGC
3121 GCGGGGAGCT GATGGTCCCC CGAGGGGTTT CTAGAGCCGC TGCCGCTACC AGCCGCTCCC
3181 CCTCGGGGTA GACATCGAAA GGAACCACAT GGCCGACCTC GGCAACCCAC TCGACCTCGA

```

3241 GATGCTCTGC CTGGTCACAG GCGGGGACTT CCGCTGGACC ATCGATTACC CGTGGGGTCC
 3301 GGGAGAGCTG TTCCTCGPAC TEGAGACCGG CGGCGAACAC AACGCGCTGC ATCAGGTCTA
 3361 TGTCACCGGG GCGACCGGAG GCACGTACAC GCTGAACGTC AACGGCACCA ACACCCCGGC
 3421 CATCGACTAC AACGACGTGT CCGGAGAATCC GCAGGGGGCTG GCAGGCGACA TCCAAGACGC
 3481 TCTGGACGCA GCCGTCCGAG CCGGAAACGC TGTCGTGCAT CCGGTCTCGC TGTTCCCTGC
 3541 GTGGACACTG AACTTCAACC TCAACGCCAG CAAGCCGCTC ACCGAGCAGT TGGTCAACAC
 3601 GATCAACAAG GCCGCGAACG ACTTCTTCGA CACGTTCGAC CAACTACTTG GGGTCGACGT
 3661 GGAGATGACG GTCACCGACA CCCTGAACCT CAAGCTCAAG GTGACCTCGC GCGCTCGTT
 3721 CGATGAGGTC GGTGTCTGTA CGTTCCGCGT CGACGTGACC AGCCAGGCAG TCATCAACTT
 3781 CTCAACTCC GTCGCCGAAC TCACCGGAGC GGTGAACACC GTCAACGTCG ACTTCTACTG
 3841 GAACCGGACG TATGACATCG AGTTCAACCG ATCCCTTGGG CTGCAGCCGA TTCCGGCTAC
 3901 TACAGCCGAC ATCACCACC TGCGGGGTAC CAGCAAGGCC GTCTCAGTCA CCGTGGTCTGA
 3961 GCCAGGAAAG AAGAGGCTGA CCATCTGGCC GTTCACGGTC AACGGTGAAG CCGCAACCAT
 4021 CAAGGTCGAG TCCGAAGAGG CCGACAAGAT CCCCACCGC TGCCGCTGGC AGTTGGTTCA
 4661 CATGCCGACC GCGGAGGCGA CCGGCGGCGA TGCAAAGCAG CTCGGCCGCG TTTACCGACA
 4141 GCCGAGGTAA CACCGCACCC ATCAGAGATG GTGGGCCAGA CCGCCTTCGG CCGCTCCCT
 4201 GACGTGTAGC TCAATGGCAG AGCGCCCGAC TGTTAATCGG GTGGTTGAAG GTTCGAGTCC
 4261 TTCCATGTCA GCGAGGGCTG AACCGGACCC GTGTCCGGTG TAGGCACTTT CCGCAGCGCG
 4321 TTCCCAAGAG CGTGGGGAGC CCCTGCCCTG TACACGTAGC TCAATTGGTA GAGCAGCGGT
 4381 CTCCAAAGCC CCGCGTCCA GGTTCGACTC CTGCGTGTGA TGCAACACC CCGTACTCCT
 4441 GCTAGCGGAG TGTTCCGCTT TCGGGCCTGG GGTCTTTTTC CCCGTTCTG TAATCGGTAA
 4501 GACACCCGCG TCTGGACCGG GCAATTGAGG TTCGAGTCCT TGCGGGGGAG CCAACTTGAC
 4561 ATCCACCCGA AAGGAACAC ATGACCTTCA CAGTCACCCG CGAGAGAGCG CAGTGGGTCC
 4621 ACACATGGC CCGCGTCCG GACGCTTCC CCTACCGCTA CCGCGGGCGG TTCACCAACA
 4481 ACCCGAGGGT GTCGACTGAC TGCTCTGGCC TGGTGCTGCA GACCGGGGCT TGGTATGGAG
 4741 GTCCGACCGA CTGGGTCTGA AACCGTTACG GCTCAACCGA ATCGTTCCGG CTCGACCACA
 4801 AGATCGTCTA CGACCTAGGG TTCAAGCGGA TGCCCCGAGG CCGGCCAGCG GCCTTGCCGA
 4861 TCAAGCCGGT GATGCTCGTC GGGCTCCAGC ACGGAGGCGG CCGGGTCTAC TCGCACCCG
 4921 CTTGCACGTT GATGACGATG GACCAACCCG GTGGCCCGGT CAAGATGTCC GACCGAGCGG
 4981 TCGACTGGGA GTCCACCGG AACCGCAACG GCGTAGGCGT CGAACTTTAC GAGGGCGCAC
 5041 GGGCATGGAA CGACCTCTG TTCCATGACT TTTGGTACCT GGACGAGTC CTCGAAGACG
 5101 AAGGAGACGA TGACGAATTG GCTGACCCAG TTCTAGGGAA GATGATCCCG GAGATCCACG
 5221 CGTGCTGTT CAATCAGACC GCGTCTGACCA GCGATCTGGC GACCCCTGGT GAAGGCGCTA
 5281 TCTGGCAGCT ACACCAGAAG ATCCACTCGA TTGACGGCAT GCTCCACCCG ATCCACGCTG
 5341 AGCGGCGCGC TCGCGCAGGC GATCTCGGTG AGCTGCACCG AATCGTGTG GCGCGGAAGG
 5401 GCTTGGGCGT GAAGCGCGAC GAGGTGACCA AGCGGGTCTA CCAGAGCATC CTCGCCGACA
 5461 TCGAGCGGGA CAACCCGAA GTACTTACG GATACATCGC AGAAAGAGGT GGCCTATGAG
 5521 CCCCAGATC CGACAGACCA TCTACCTGCT CCGCACCGCC GCCCGCGAT TGCTGGGCAT
 5581 CGTCTGATC TGGGGCGGGC TCGACGCTGA GTCCGCGGCT GACCTCGGTG ACATCATTGC
 5641 GGGCGTCGTG TCGATACTAG TCTCCGGTGC GCCGGCCGTA GCGGCAGGCA CCGTACCGAG
 5701 CCAGCGCAAG GACGGCACGT TGTCACCAAG CCCGGTGGAT CAGGTACCA AGGGCGTCGA
 5761 GCAGGTGCTC GCGGCCAGGC AGAGTGCCGA GGCTGAAGTC GCGAAGGTCA AGCAGGCGCT
 5821 GGAGACCGCC GTCAGCGGTT CTCTCCCCCA GCTCGGCCCG CTGGCCACGC AGATCCTCAA
 5881 CGTGGCTGAC GACACCGTCT GCGCTCCATG AGCAAGCCCT GGCTGTTAC CGTCCACGGC
 5941 ACAGGCCAGC CCGACCCGCT CCGGCTGGT CTGCTGCGG ATACCGCACG GGACGTACTT
 6001 GACATCTACC GGTGGCAGCC CATCGGCAAC TACCCGGCAG CGGCGTTCCT GATGTGGCCG
 6061 TCGGTGAAA AGGGTGTGCG TGAGCTGATC CTGCAATCG AGCTGAAGCT GGACGAGAT
 6121 CCGTAGCGCG ACTTCGCGCT GGCGGCTAC TCGCAGGGAG CCATCGTGGT GGGCCAGGTG
 6181 CTCAAGCACC ACATCATCAA CCGAGAGGT CGACTGCACC GGTTCCTGCA CCGGCTCAGG
 6241 AAGGTCATCT TCTGGGGTAA TCCGATGCGG CAGAAGGGCT TTGCCCACAC CGACGAGTGG
 6301 ATTACACAGG TCGCTGCCTC GGACACGATG GGATCCTCG AGGACCGACT GGAGAACCTC
 6361 GAGCAGTACG GCTTTGAGGT CCGGACTAC GCGCACGACG GCGACATGTA CGCCTCCATC
 6421 AAGGAGGACG ACATGCACGA GTACGAGGTG GCCATTGGCC GAATCGTGT GAGCGCTAGG
 6481 CGATTATCG GAGGTAAGGA CTCGTCATC GCCAGCTCA TCGAGCTTGG ACAGCTCCG
 6541 ATCTGGGAGG GAATCGCGAT GGCCAGAGCC ATCATCGACG CCCTCACGTT CTTCCGCAAG
 6601 TCGACCCAAG GCCCGAGCTG GCCGATTTG TACAACCGCT TCCGCGCGGT CGAGTTCCTA
 6661 CGACGAATCT GAGAAAGGAG GCGGGGTGAG CCTCAACAAC CACCACCGG AGCTTGCCCC
 6721 GTCTCCCCCT CACATCATCG GCGGCTCTG GCGAAGACG TCCGAGTGGT GTGCTATCT
 6781 GCCTGAGAAG ACCCTCGGCT GGGGAGTCTT GAAGTGGCTC TCCGAGTACG TGAATACCCC
 TGGCGGGCAT GACGATCCGA ACCGTCTGGC GACGTTGATC GCGCTCTCCG AGGCAGGTCT

6841 TCTCGACAAC GAGAACATGT TCATCCCCAC CGACGAGCAG GTACGCCTGG TCCTCTGGTG
6901 GTACGCACTA GATGACCAGG GCGAGTACAT CTACCGGAG GGCCTGATCC GCCGGCTCAA
6961 GGGCTGGGGC AAGGATCCGT TCACCGCCGC GCTCTGCTTG GCGGAACCTCT GTGGCCCCGT
7021 AGCCTTTTCA CACTTCGACG CCGACGGTAA CCCGGTCGGC AAGCCGCGTT CAGCCGCGTG
7081 GATCACCCTC GCGGCCGTCA GCCAGGACCA GACGAAGAAC ACCTTCTCGC TGTTCCCGGT
7141 GATGATCAGC AAGAAGCTGA AGGCCGAGTA CGGCCTGGAC GTGAACCGCT TCATCATCTA
7201 CTCCGCAGCC GGTGGCCGTA TTGAGGCAGC GACCTCGAGC CCCGCGTCGA TGGAGGGTAA
7261 CCGCCCCGACG TTCGTCGTCC AGAACGAGAC GCAGTGGTGG GGCCAAGGCC CCGACGGCAA
7321 GGTCAATGAA GGCCACGCGA TGGCAGAGGT CATCGAAGGC AACATGACCA AGGTTCGAGGG
7381 CTCCCGCACC CTGTCGATCT GCAACGCCCA CATCCCCGGC ACCGAGACGG TCGCCGAGAA
7441 GGCATGGGAC GAGTACCAGA AGGTCCAGGC AGGCGACTCT GTCGACACCG GGATGATGTA
7501 CGACGCGCTG GAAGCGCCGG CCGACACCCC GGTCTCCGAG ATCCCCCGCG AGAAGGAGGA
7561 TCCCGAGGGA TTCGAGAAGG GCATCGAGAA GCTCCCGGAG GGCCTGCTCA TCGCCCGAGG
7621 CGACTCCACC TGGCTGCCGA TAGACGACAT CATCAAGTCG ATTCTGTCTA CCAAGAACCC
7681 GATCACCAGG TCGCGGCGCA AGTTCTTGA TCAGGTAAAC GCCGCTGAGG ACTCGTGGCT
7741 CTCACCGCAG GAATGGAACC GGTGCCAGGT CGACCTGGCC AAGTACCTGG ATAAGCACGG
7801 CAGGGAGTTC GCTCCGCTGC AGCGCGGTGA CCGGATCACC CTCGGGTTCG ACGGGTCGAA
7861 GTCCAACGAC TGGACCGCGC TCGTCGGCTG CCGTGTGAGC GACGGCCCTGC TGTTCTCAT
7921 CGACATCTGG GATCCCCAGA AGTACGGCGG GGAGGTTCCG CGCGAAGACG TTGACGCCAA
7981 GGTCCATTCG GCGTTTCGCC ACTACGACGT GGTGGCGTTC CGCGCCGACG TGAAGGAGTT
8041 CGAGGCGTAC GTCGACCACT GGGGCGGGAC CTACAAGAA AGCTCAAGG TCAACGCCAG
8101 CCCGAACAAC CCGGTGGCGT TCGACATGCG CCGACAGCAG AAGAGGTTCC CGTTCGACTG
8161 CGAGCGACTC GAGGACCGCG TCCTTGAGGG CGAGGTCTGG CACGACGSCA ATCCCGTTCT
8221 GCGCCAACAC GTTCTGAACG CCAACGACA CCCACGAAC TACGACGCCA TCGGATTCG
8281 CAAGGTCAGC AAGGACTCCA GCAAGAAAT CGACGCTGCA GTCTGCGCTG TCCTCGCGTT
8341 CCGGGCGAGA CAGGACTACC TCATGAGCAA GAAGGCCCGT AGCGGCCGGG TGGTGATGGT
8401 TCGATGACAG CACCGCTCCC CGGTATGGAG GAGATCGAAG ACCCCGCACT CGTACGAGAA
8461 GAGATGATCT CGGCTTCGA GGATGCTTCC AAGGATCTCG CCAGCAACAC CAGCTACTAC
8521 GACGCTGAGC GCGGCGCAGA GGCCATCGGC GTCACCGTCC CGAGAGAGAT CGAGCAACTG
8581 CTGGCTCAGC TCGGATACCC CAGGCTCTAC GTCGACTCAG TCGCCGAGCG CCAGGCCGTC
8641 GAGGGTTTCC GCCTCGGCGA TGCCGACGAG GCTGACGAAG AGCTGTGGCA GTGGTGGCAG
8701 GCCAACAACC TCGACATCGA GGCACCACTG GGCTACACCG ACGCTTACCT TCACGGCCGG
8761 TCGTTTCATCA CGGATCAGCA GCCAGACCCG CAGCTCGACC TGGGTTGGGA TCAGAAGCTC
8821 CCGATCATCC GCGTCGAGCC GCCACCCGGA ATGCACGCCG AGATCGACCC CCGGATCAAC
8881 CCGGTGTCCA AGGCCATCCG AGTCGCATAT GACAAGGAGG GCAACGAGAT TCAGGCTGCC
8941 ACGCTGTACA CGCCGATGGA GACCATCGGC TGGTTCCGGG CTGACGGTGA GTGGGCTGAG
9001 TGGTTCACCG TCCCGGACGG TCTGGGCGTC GTTCCCGTTG TGCCGCTTCC GAACCGGACC
9061 CGGCTCTCGG ACCTGTAGGG CAGCAGTGAG ATCAGGCGCG AGCTTCGGTC GATGACGAGC
9121 GCGGCGGCGC GCATCCTCAT GTTGATGCAG GCGACCGCGG AGCTGATGGG TGTCCTCCAG
9181 CGCCTGATCT TCGGATCAA GCCGGAAGAG ATCGGCGTCC ACTCCGAGAC CGGCCAGACG
9241 CTGTTTCGATG CGTACCTGGC CCGGATCCTG GCGTTTCGAGG ACGCTGAGGG CAAGATCCAG
9301 CAGTTCTCTG CAGCCGAGCT GGCCPACTTC ACCAACGCGC TCGATCAGAT CGCCAAACAG
9361 GTCGCTGCGT ACAGGGGATT GCCTCCCCAG TACCTGAGTA CCGCCGACAG CAATCCGGCC
9421 TCCGCTGAGG CGATCAGGGC CGCTGAGAGC CGACTCATCA AGAAGGTCGA GCGGAAGAAC
9481 CTGATGTTTCG GCGGCGCATG GGAAGAGGCC ATGCGGATCG CCTACCGGAT CATGAAGGGC
9541 GGCGACGTTT CCCCAGACAT GCTCCGCATG GAGACCGTCT GGCGAGACCC GAGCACTCCC
9601 ACCTACGCGG CCAAGGCCGA CGCAGCCACG AAGCTGTACG GCAACGGCCA GGGTGTCTATC
9661 CCGCGTGAAC GTGCTCGCAT CGACATGGGC TACTCCGTCA AGGAGCGCGA AGAGATGCGC
9721 CGATGGGACG AGGAAGAGGC CGCAATGGGT CTCGGCCTGT TGGGACGAT GGTTCGACGCC
9781 GACCCGACGG TCCCAGGCTC CCCGAGCCCC ACGGCACCGC CGAAGCCACA GCCGGCCATC
9841 GAGTCGTCTG GTGGTGATGC GTGACCGCAG AGGAGTACGC GGCGGCTCAA GCCCGATCA
9901 CTGCGGGTCT TGCCACATAC GTCCAGAGGT TCGCTTCGCT CTTCGTCTGGT CCAGCTCTCG
9961 CTGTAGGTGA GTGGCTGCGA CTGCTGCAGG TGCTGTTCCC CGAAATCCAA CCGCGGTATG
10021 CAGATGCTGC CGCCTTGCGC AGGGACTTCT ACGACTCCCA ACGCGCACTA CACCAACCCAG
10081 AGCTGCCCCG GAACGAGAGG TTCTGCGGAG AGCTTCGGTG GGAGTGTTTC GTCCAGAAC
10141 TGGAGCCCGC TCGAAAAGAG ATGTCGAGG CCGACTCTCC CGCGAGTCCG ACCTCTAAGT
10201 TGGCTCTGGC CGCAGTTCGC GAAGTGGAGA TGGCAGTACG CCGACAGATC ATCGGCGCTG
10261 TCAAGAACGA TCCGGCCCCG CAGATCGTGC AGGGCTGGGC GAGGGTCCGC ACCGGGCGCG
10321 AAACATGCGC CTGGTGTCTG ATGCTCATCT CACGGGGTGC CGAGCTGAAT CACAAGGGCA
10381 ACTTCGCCTA CAGCTCAGCG GAAGCCGACG GGCTCAACCT CGATGACGAG ACCGTGATCG

10441	ACCTCTGGAA	CGAGTCCGGT	CACGACCTTG	AGAAGTTCCG	CGAGGAGACC	AGAGAGGACT
10501	TCGAGAAGTG	GCACGCAGGG	TGCGACTGTC	TGGTGGTCCC	GGTCTTCGAT	GTGCAGAACT
10561	GGCCCCGAAG	AGACGCTGCC	CTACGGGGCG	AGCAACTTTG	GATCGAAGCC	AGCGACGAAG
10621	CTGACGACCT	CATTGCGTCA	GGCAAGGCCC	GCTCCAAGAA	CAAGAACACG	GAGACGCTCA
10681	ACGCGCTCCG	ACGCCGCCTA	GCACGCGGCG	AAATCACCAT	GTCCAACCTAC	GCCCTCGCTG
10741	CGTAGTCCCT	CGAACCCAG	GTGGGTCTCT	TCAACATGCC	CAGGAGGCGA	AAACACATGT
10801	CCGACAACCC	CACTCCCGAG	AGCACCCAG	AGGCCGAGAC	CCCGGAGGTC	GAGAAGCCGA
10861	TGGAACCGCA	GGGCAAGGTC	TTCGATGAAG	CGTACGTTCA	GTCGCTTCGC	CAGGAGGCTG
10921	CAGCCGCTCG	GGTGGCGAAG	AAGGACGCGC	TAGAAGCGGC	AGAGGCTCGA	GTGAAGGCCG
10981	ACTACGAGGC	CAAGCTCGCT	GAGCGCGACA	CCGCTTACAC	CGAACTGCAG	AACCAAGTTGG
11041	GACAGGCGTG	GATTGAGCTG	GAGAAGGTCT	ACCTCTCTCT	CGACGCCAAG	GTGCCCAACG
11101	ACAAGGTTCC	GGCGTTTGTC	GAGATCCTCG	AAGGCAACGA	CAGGGACAGC	ATCGCTGAGT
11161	CAGTGAAGTC	CCGTCTGGAG	CTGGTCCGCG	GATTCCGGCA	CAAGACCCCG	AGTCCTGCGT
11221	TCGACCCGTC	TCAGGGTCGC	GGCGGTAAGC	CGCCGATCCC	GCTGAACGGT	GACCCGATCC
11281	TCGAGGCCAT	CAAGGCCGCT	GTCGGGATCA	AGAAAGTAACC	CACCCAACAG	ATCTCAAGGA
11341	GAGATAAACA	ATGGCAGTCA	ACCCTGACCG	CACCACGCGC	TTCTCTGGCG	TGAACGACCC
11401	CAAGGTCGCG	CAGACCGGCG	ACTCGATGTT	CGAGGGGTAC	CTCGAGCCCG	AGCAGGCCCA
11461	GGACTACTTC	GCCGAAGCGG	AGAAGATCTC	CATCGTCCAG	CAGTTCGCCC	AGAAGATCCC
11521	GATGGGCACG	ACCGGCCAGA	AGATCCCGCA	CTGGACCGGC	GACGTGAGTG	GCTCGTGGAT
11581	CGGTGAAGGC	GACATGAAGC	CCATCACCAG	GGGCAACATG	ACCTCGCAGA	CCATCGCCCC
11641	CCACAAGATC	GCGACGATCT	TCGTGGCCTC	GGCGGAAAGC	GTCCGTGCGA	ACCCGGGCCAA
11701	TTACCTGGGC	ACCATGCGGA	CCAAGGTGCG	GACCGCCTTC	GCGATGGCGT	TCGACAACGC
11761	CGCCATCAAC	GGCAGCGACA	GCCCGTCCCG	GACCTTCCTA	CCGCAGACCA	CCAAGGAGGT
11821	CTCCCTGGTG	GACCGGGAAG	GCACCGGCTC	CAACGCGGAC	CTCACCGTCT	ACGACGCGGT
11881	CGCCCTCAAC	GCCCTGTCCG	TGTTGGTCAA	TGCCGGCAAG	AAGTGGACCC	ACACTCTGCT
11941	GGACCAACATC	ACCGAGCCGA	TCCTCAACGG	CGCGAAGGAC	AAGAGCGGTC	GCCCCGCTGTT
12001	GATCGAAGTG	ACCTACACCG	AGGAGAACAG	CCCGTTCCGC	CTCGGTCCGA	TTGTGGCCCC
12061	TCCGACCAAC	CTGAGCGACC	ACGTCCGCTC	GGGCACGGTC	GTCGGCTACC	AGGGTGACTT
12121	CCGCCAGCTC	GTCTGGGGCC	AGGTCCGCGG	CCTGTCTCTC	GACGTGACGG	ATCAGGCCGAC
12181	TGTGAAGCTG	GGCAGCCCCC	AGGCTCCGAA	CTTGTCTCTG	CTGTGGCAGC	ACAACCTCGT
12241	CGCAGTCCGA	GTGAGGCGCG	AGTACGCCCT	CCACTGCAAC	GACAAGGACG	CGTTCTGTCAG
12301	GCTTACGAAC	GTGGACGCCA	CCGAAGCCTG	ATCCAGGCTT	GACATCCACC	GGGAGGGGGC
12361	TCCTTCGGGA	GCCCTCTCTC	GATGTGAGAG	AGGAAGGACC	ACATCGGAAT	CCAGTCCACC
12421	CTCAACGCGG	GTTCGCGCGA	GGTTCCCGAG	GAGTTCGCCA	AGCAGTTGAT	CGCCACTGGC
12481	GGCTGGGAAG	TGCCCCGGAA	ACCGCGCAAC	ACCAAGACCA	AGACCGCTCC	TGAGGAGCCC
12541	AAGAAGGAGG	AGTAACCCGT	GGCCTACGCG	ACCGCCGAAG	ACGTTGTGAC	GTTGTGGGGC
12601	AAGGAGGCTG	AGCCCCAAGT	GATGGCGCTG	ATCGAGCGCC	GGTCCAGCA	GATCGAGCGC
12661	ATGATCAAGC	GCCCGATCCC	CGACCTGGAG	GTGAAAGCGG	CTGCGTCCGG	GACGTTCCGG
12721	GCCGATCTGA	TCGACATCGA	AGCTGATGCT	GTTCTGCGCC	TCGTGCGTAA	CCCGGAGGGC
12781	TACCTCTCGG	AGACCGACGG	TGCGTACACC	TATCAGCTCC	AGGCCGACCT	GTCGCAAGGC
12841	AAGCTCACCA	TCCTCGATGA	GGAGTGGGAG	ATCCTCGGGG	TCAACTCCCA	GAAGCGCATG
12901	GCGGTCATCG	TCCCGAACGT	GGTGATCGCG	ACGTGAGCGC	GAGCGACCGA	CACCGCGCCC
12961	CGATTGTCTA	TCCGCTGGC	ACTCAGGCGG	TTACGCCGGA	TCCGGTCAAC	CGGTTTGACT
13021	GCGATCACGA	AGCTGATCCT	CCGGTGTGCC	GGTGGCTCA	CGACTGGCGC	ATCGAGTGGG
13081	GAAACGTCAA	GAAGGCCACC	GCCAGATCAC	GGTCCGGCGT	GCTCTGATGA	GCCTCCTCGA
13141	CACCGGTGCC	CGGTACCAGA	CTGTCATCGT	CTACCCCGAA	GAGATGGTCA	TCGACTCCGA
13201	TGGCAACAAG	CGGACCAGGC	CGTCGAATAC	CGGCATCCCG	GCCATCCGAC	GGTTCAGGT
13261	AGCCAACCAG	TCTGGTACGT	CGGCACGACG	TGCTGAGCAG	GACAACGAGG	GGTTCGAGAC
13321	CGAGAAGGTC	TACCGGATGC	GGTTTCCCGG	CTCGTTCACC	AAGGAGCAGG	GCATCCTCGG
13381	GGCCCCAGTC	CAGATCGAGT	GGCGAGACCA	GCGGTGGGCG	CTCTTCGGAG	ACGCCACCGT
13441	CTACGACTCA	TCCCTTGGCT	TGGCGGGGCT	CGACTACAGG	ATCAAGAGGT	ACTGATGGCC
13501	AAGGTCTACG	CGAACGCGAA	CAAGGTCCGC	GCCCCGTACG	TGCAAGACGAG	AGACGCCGTC
13561	CGAGACGAGC	GGAAACAAGT	CACCCGTGCA	GCCAAAGCCA	ATCTGGCGCG	GCAGAACTCG
13621	ACCACCCGCA	TCACCGACGA	GGGCTACTTC	CCGGCCACCA	TCACCGAGCA	AGACGGCGAT
13681	GTCGACTTCC	ACACGATCCT	CAACGCGCCC	AACGCGTTGG	CGCTTGAGTT	CGGCCACGCG
13741	CCGTCTGGCT	TCTTCGCTGG	CACCGACAGG	AAACCAACCG	AGGCCACTTA	CATCCTCACC
13801	CGAGCCGCCA	TCGGCGGCAC	CGTCTCATAA	GGAGGTCACA	TGGCGCGAAT	GCCTCGCGTC
13861	CAGGCAGTAG	GGGCCCCGAT	CCTCCGGTCA	GACCCCGGAC	TGGAGGGAGT	GACGGTACAG
13921	ACATGGGTTT	CAGACGTGGA	CTTCCGAGAG	TTCCCGATGA	TCAACCTCCG	CCGCATAGGC
13981	GGGACGAGGA	ACCCCAACGC	ACCGACGCTG	CACACGCTGC	CGGTGGTCCA	AATGACCGCC

14041	TACACCAGAG	ACGGTCTCAT	GGAGACTGAG	GAGCTGTACG	AGACCGCGCT	AGAGGTTCTC
14101	TACGAGGCGG	TGGAGAAGGG	AACACAAACT	CCCGCAGGGT	ATTTGACCTC	CATCTTCGAG
14161	ACGATGGGCG	CCACTCAGTT	CAGCTCCCTC	TACCAGGACT	CCTGGCGCAT	CCAGGGTCTG
14221	ATCAGGCTCG	GCGTCCGEAG	ACCGAGAACC	ACCCTCTAAC	CGAAAGGTAA	AGCCACATGG
14281	CTGAAAACGA	CGACGCAGTG	TTGACTGCGG	CGGTCCGGTA	CGGTACGTC	GGTGCTGCAG
14341	GCACGGCTGC	TCCTACGCCG	GCCTTGCTCA	AGACCATCGA	CCTCAGCAAG	CCCGAGACCT
14401	GGACCGGTGC	TACCGGTTGG	ACGAGGCTCG	GCCACACCAG	CCGAGGCAGG	CTCCCTGAGT
14461	TCGGCTTCGA	AGCGGCGGAG	TCCGAGGTCA	AGGGCTCCTG	GCAGAAGAAG	AAGCTCCGCG
14521	AGATCACCAC	CGAGGATCCG	ATCGACTACG	TCACGGTCCT	ACTGCACCAG	TTCGATGAGC
14581	AGTCGGTGGG	TCTGTACTAC	GGCCCCAACG	CCTCTGAGAC	TCCTGGTGTG	TTCGGTGTGA
14641	AGACCGGCCA	GACCAACGAG	AAGGCCGTGC	TGGTCGTGAT	CGAAGACGGC	GACATGCGCC
14701	TGGGGGATCA	CGCCCAACAG	GCTGGAGTTC	GCCGCGACGA	CGCGATTGAG	CTGCCCATCG
14761	ATGACCTGGC	TGCGCTGCCC	GTCCGGTTCA	CCTACCTGGA	CCACGAAGAC	GAGCTGCCGT
14821	TETCCTGGAT	CAACGAAGAC	CTCTTCAACG	TGCCCGAGGT	TCCCCAGGGG	TGATCCCAAC
14881	TTGACAGCCA	CCCCTGCTGC	TACCCCGGAG	GGGGAGGTTT	CCTTGGCAGG	CCTGGCCTCC
14941	CECTGCTCCC	GCCACTEACA	GACCCGCGCA	CACTGAAAGG	TTCCGCTATG	CAAACTGATT
15001	CACCATGAC	GCATTCCGCG	AAGAGGTCAA	GAAGAAGTAC	GCTCCGGTCC	TCATCGGCCT
15061	GTCCGACGAT	GTGACCGTCG	AGCTGAAGCC	GCTGCTGAAG	CTGGGCCAGA	AGGCCCGCGA
15121	AGCGGTGGTC	GAGGTGTTCA	AGGAGTTCGC	GGACATCCCC	GACCTCGAAG	AGGACGACGA
15181	CGACGAGTTG	GTCGATGAGT	ACTGCTGCTA	GGTCTGCGAC	ATCATCGCCA	AGGCGTTCCG
15241	GCTGATCGCC	ACGAAGCCCA	AGAAGCTGAT	CGCCGECTTG	GACGAGGAGC	CGGATCCCCG
15301	TATCCGCGCA	GAGCTGTATG	CAGCGTACT	CAACACCTGG	AAGCGAGAGA	CGCAACTGGG
15361	GGAAAGCCCG	CCCTCGCCGA	GCTGATGAGC	AAGTTCGCGG	GGGCGATCCT	CGCAGACCTG
15421	CTCCAGTACT	ACCGGGTAGA	CCTGCGGAGC	CTGTTCCGCG	ACGAGGATCC	GCTTTCGCGG
15481	AGATTCTGTT	TGTCCTGGGT	GCTGCTGCTT	CCCAAAGACG	CGCGCTTCTA	CGCAGAACGT
15541	CGTGGTGGGC	AGCAGTACCG	GGGTGCGACC	GAGGACCGCT	ALCGCTCTCG	GGCATCTAC
15601	GACGCCATCC	AGGCGGGCAA	CCACATCTTG	CTGCTGGCGA	ATCGTGATCC	GAAGAAGCCA
15661	AAGCCCAAGG	GACCCAAGTC	ATACCCGCGT	CCCGACGACC	TAGAGAAGAC	CACACCGAAG
15721	CCGGGTTTGT	TCGCGGCAAT	GGTCGTGCGA	GCGAAGAAGG	CGGCTCGAGA	GAGAAGGGAA
15781	AGGGAGGAGG	AGAGTGCCGA	ATAGTGCTGG	CGTAGAAGTC	GCCCGGATCT	CGGTCAAGGT
15841	CAGCCCGAAC	ACCAAGGAGT	TCCGCCGGGA	ACTCAAGACC	GAAGTTCGAG	AGATCGAGCG
15901	GGAGGTTAAG	GGCGATGTCT	AGATCAACGG	TCATCTCGAT	GCGGCCCAGG	CCAAGGCCGA
15961	CTTCAAGCGC	ATGATGATGC	AGCTCAAGAC	CGAAGCTGCC	AAGGGCGTTC	ACGTCCCCGT
16021	GCACGTAAAC	GTGACAAGA	AGAGCAAGAA	GGGAGGTCTC	CTGAGCTGCT	TCCCTCCGCG
16081	CAGCCGGGGG	CTCGGAGATC	TAGGCGATGA	CGCCGAGAAG	CGCTGCTCTC	AAGTACAACA
16141	CCTTGGCAAG	TCGTCTCTGG	GCCTCACACG	AGCCGCCTGG	ATAGGCGTAG	GCATCGTCGC
16201	CGTAGCAGCT	CCGCTGGTCT	GCATCGTGCC	CGGTCTGCTG	GCCGGTCTGC	CGTCCGTGCT
16261	GTCTGCGTTC	GGAGCCGGCG	CTGGCGTAGT	CGCGCTCGGC	ATGGACGGCA	TCAAGGCAGC
16321	CGCCTCGACG	CTGGCCCGGA	CGTCAAGGCC	GGTCTCTCCT	CTGCTCTCCA	CGACGTTCCA
16381	GCAGGGACTC	ACCCCGGTGT	TCCAGCAGCT	CGGCCCGATG	CTGACCGCGA	TCACCCCAAC
16441	CCTGCAGAAC	GTGCGCTCGG	GCCTCGTGAA	CATGGCCGGG	TCGATCACC	ACGTGATCAC
16501	CCAGGCTCCT	GGTCTGCAGC	AGATCCAGAA	CATCCTCACC	AAGACCGGAG	AGTCTCTCAC
16561	GGGCTCTGGC	CCTGTGCTCG	CTACCGGCAC	GCAGGCGTTC	CTGACGCTGT	CCAACGCCGG
16621	CGCGAACTCG	TTCGGCACGC	TCCTGGCTCC	CCTGCAGGAG	TTCAACACCG	GCTTCAACGA
16681	CATGGTCAAC	CGAGTCAAGT	CCAACGGCGT	GTTTCGAGGT	GCCATGCAAG	GGCTTTCGCA
16741	GACGCTGGGC	AGCGTCTCTA	ACCTGTTCAA	CCGGCTCATG	GAGTCCGGTC	TGCAGGCGAT
16801	GGGACAGCTC	GGCGGTCCGC	TGTCGACGTT	CATCAACGGG	TTCCGAGATC	TCTTCGTCTC
16861	GCTGATGCCG	GCGCTGACTT	CGGTCTCTGG	TCTGATCGGC	AACGTCTCTG	GGACGCTGGG
16921	CACACAGCTC	GCTCCCATCG	TCACGGCGCT	CACGCCGGCC	TTCCAGACGC	TGGCGAGCAC
16981	GCTCGGCACG	ATGCTCACC	GAGCCCTCCA	AGCTCTGGGT	CCGATCCTGA	CTCAGGTCCG
17041	TACGTTGATC	GGCAGCAGCG	TGAACACGGC	GCTGCAGGCT	CTCCAGCCGA	TGCTGCCGTC
17101	GCTCATGCGA	AGCTTCCAGC	AGATCTCCGA	CGTACTGGTG	ACCACTCTGG	CCCCGCACAT
17161	CCCGGCGCTG	GCGACGGCCC	TCGGCCAGGT	CGCAGGCGCG	GTGCTGCAGC	TCGCTCCGAC
17221	GATCATCTCG	ACGTTGGTTC	CGGCGTTTCT	TCAGTTGGTC	CCAAAGGTCT	CTGAGCTAGT
17281	TCCGACCATC	GTCAACCTGG	TCCAGTCTGT	CGCCAACCTG	ATGCGGTTGT	TTCTGCCCTT
17341	GGCGCAGGCT	TTGGTCAGCG	TTGCTGGGCG	GGTGATTGAT	GTGGGTCTCT	CCATCGGCGG
17401	CGCGCTCATC	GGCGCGCTGG	CGAACCTCAC	GGAGATCATC	TCCAACGTCA	TCAAGAAGGT
17461	GTCCGAGTGG	GTCAGCAGCT	TCTCCAGCGG	AGCCAGCAG	ATCGCTGCGA	AGGCAGCGGA
17521	ACTGCCGGGG	ATGATCCAGT	CGGCTCTCCG	CAACCTGATG	GCCATCGGGC	TGCAGGCGGG
17581	TAAGGATCTC	GTCCAGGGCC	TGATCAACGG	CATCGGCGGG	ATGGTCAGCG	CAGCGGTCAC

17541 CAAGGCCAAG GAGCTGGCGT CCAGCGTGGC TGGTGCAGTG AAGGGCTTCC TGGGCATCGA
17701 GTCCCCGTCC AAGTTGTTCA CCGAGTACGG CCAGTTCACC GCGGAGGGAT TCGGCAACGG
17761 CATGGAGGCA GGGTTCAAGC CCGTCAATCGA ACGGGCCAAG GATCTCGCGG CTGAGCTGTC
17821 CAGGGCGGATC GAGTCGGGCA CCGACCCCTC CCGGATTCTC GCTGGGCTGG ATCAGAATGA
17881 GCTGAAGCAG ATGCTGGCGG CTCTCGAAGA GGAGCGCAAG CGACTCAAGG TCGAGAAGAA
17941 CGGTATCCCC AAGGGAGACA AGGCAGGCCG AGAGGCGCTG CAGAACCAGC TCGACCAGAT
18001 CCAGSCGCAG AAGGACATCC TGTCCTACCA GCGTGACCGC ATCAAGAACG AGTCTGAGTA
18061 CCGCGACATG GCGGCCAAG ACCCGTTGGT GAAGGCAGCC TCCGGGCTGA TGAGCGCACC
18121 GGTCGACTTC GCGAAAGCGA CTGGCAAGCA GTTCCTTTTCG GACATCGGCA TCAGCGGAGA
18181 TGGGTTTCATC TCCAAGGCCA TCACCGAGGG CATCCAGTAC ATCTTCCAGA TCGGCTCTGT
18241 CGATGAGGCG CTGTCCATCA AGGACCCGGA GGAGTCGAAG AACCGCGTGT CCGTCTGTTGG
18301 CCGCTGACTT GACATCCACC AGGAGGTAAG CATTGATCAC CGACACCATC GTTGAACTCG
18361 AGGGTGTCAA TGGTGAGCGT TTCAACTTGA CGACCGGTGA CCAGGGTGTG TACCTGGCCA
18421 CAGACGTGGA GGGTTGTTTC TACGACCCCTC CCGTCAAGGT CGTTGTTGAA GAGCCGGGGA
18481 TCTACCCCGG CGCTCGCTAC TTGTCCCAAG GAGCCCTGAA GCGAGACATC GTCTTTGGGG
18541 TCGTCACTCT CAACGACCGG AAGCAGGCGC CCGGCTCCTG ECTGTCCGCA GACTCCGAGT
18601 GCGGCAAGGC GTGGGCGTTC AACCAGCACT GCAAGCTCTA CGTCAACCACC CCGGACTCCG
18661 GTACCCGCTA CCTGAAGCTG GCGCTGTTCG AGTCCCCCAC CGTCAAGATG GACACCGACC
18721 CAAGAGGTA ACCCCTTGAG GTCACGGTGA TGTCTGTGAT CCGGTACGAC CCGTCTGGT
18781 ACGAGGACGA CAAGGCTTTC TCGGCCAAGA CCAAGACCGA CACCCGGTTC GACCCGCTGT
18841 TCTGGACGCC GCGGTGGCGG TGGGAGGAAC TGCCCAAGGA GACGCTGCGG ATCAAGGTGG
18901 GCGGCGAGCA GGGTGGGCTA AACCACCGG ACCAGTACAT CTTCCCGAAG TGGACCGTTC
18961 CCGGCTCCAC CGAGAAAGTG CCGAATCTCC CCTGGCCGTT CCCCCGAAC GTCCCGATCC
19021 CTTGGGAGAC AGCACCGTTC ACTCAGTTCG TCATCCCGGA CTACTCGTTC GAGGATGAGG
19081 AGTTCCGCAA CCGCCGCTC AAGACGTCG GGTGATCTA CCGCGAGAAC TCGCTCATCG
19141 ACACCGACCG GCGCGAGGAG CAGATCGCTT CCGAGTCGGG CTCCCGGCTG TGGGCTCGGA
19201 TGAACGGTGT CCGGTTCCGC AACTCGATCC GCGCCTACAC CGAAGAGGCT GAGTTCGTCA
19261 TAGACGCATC GGGATGCGCT CCGGACAGG TAGTTACCTT CCGGCTCACG AGGCCGTTGT
19321 CCGGCTGCTG GGGGCTAGAG TGAGTGGTCT GACGAGCGTT CGTGAGGCCG AAGATCTCTG
19381 CCAGAAGATC CAATTGCGGC GCTGCAAGCG CGAGCAGGAA CCGCTCAAGC ATCCCGACGT
19441 AGAGCTGCGC GATGGCGACT TCCGCTGCG CCGCCTGGTC GCTGGCGAGC GGGTGCTCGA
19501 GTGGGAGTTC ATCGAGAACC AGACTGGCAC CTGCACCTTG CAGCTCTCAC TGAGCCATTA
19561 CCTGGCGAAG TGGGTGATGG ACCACCGGGG TCGAGCAAAG CGCAACGTCA TCATCAACAT
19621 CGAGAAGCAA GCGGCTCGAT GGACCGGGAT GATGGACCA TACCGGGTCA TCAAGACCGA
19681 CCGAGGGGAC GCCTACATCG AGATCGTGTG TTTGCACGAC TTCGAGCAGA CCAAGCATAT
19741 CCGGGTATGG TGCAACCCGT TCCTACGCCC CGAGCTGCAG TTCCCAAGG TGTGGATCAT
19801 CTTCCGGCGG GCGAAGTGGT GTTTGCTGGT GACACTGTTG GTCAACCTGC TCAGGCTCGA
19861 GACGAGCTTG TGGACGCTGC CTGATGACCC CACGGACATC AACGAGTGA TGGGTCCGAG
19921 CTTCAACCCA GCAAATTGGC GGAACATCGT CAAGCCGTTT CCGTTCCTGG CCGACAATC
19981 ACCGGTCACG ATGGTGTTC ACGCGTTCGG GACGTTCTAC GACACCGCCA AGAAGATCCT
20041 CGAAGACCAT CAGCTCACGC TGACGTGTCG TCGGTACATC AAGGACCGCG ACCCGCATCC
20101 GTTCGAAGAT CTCAGGGGGC TCTGGGGAAT TGATCCTGTC GAAGACCTGC TGCAGAAGAT
20161 CCGGCTCCGG GACGGCTGCG TGGTCTGGGA CATCGAGGAC AACTCAGGTT GGGGCACTCA
20221 GACCGCGTTC GCGGTTCCGT GGCTGACCGG GTTCGTCCGA GGGATGGTCC AACTGGCCGG
20281 CGACGGCCAG GTCGAGGGCG TCGATGTGTT CACCGGGGAC TACACGTTC CAGGCGAGTA
20341 CTACTCCCCC TGGTTCATGG GCACCGAGCC GATAGCACCC CACGTCGTGT TCGAAGAAGG
20401 ACCGCTGACC GGGATCAAGT CGTCGGAGTT CTCGTACTAC GAGGCCACCG ACACAGCTT
20461 CCTGGCTGGT GGACAGAGCG CACCTGGCAT CAACGAGGGC ATCTCGGCC TGGTGAACAT
20521 CCGTGGCGAC CTGCTGACCT CGTTCATCAA CAGCCAGCTC GCGCGCTCG GCGCGGTGCG
20581 TGGAGCGATT GACCTCCCGC CTCTGGGCGG TCTGCTCGAT CCGGTGTTG AGCCTCTGTA
20641 TCCGATGTG TTCGGCGCGT TCATGGAAAG TCCGACTCTG CGTGAGTGG GCATCTCGCT
20701 CCCGATCTCC GGGCTCGAGG ACATCGTCAC CCGACTGGGC GACTTCCACT ACTTCGAGAA
20761 CATGGCCGAC GGGCGCATGA AGGCGTTCAC GCTGTGAGCG TTCGAGCCA TCGCATCGCA
20821 GATCCACAAG ACGAGGGCTC GAACGACCCA CACCTCAAG GTGTCTGACG CCGCTCCGTA
20881 CATCTTCGCG CCAAAGCCCT ACGGCACTG CTGGATCGGA GATCGCGTCG GCACGTCGGT
20941 CCTCGGCTAC CCGTTCGAGC CCGTGTGTT CGTGAGCGC ATCCGCAAGG TGAAGTACCG
21001 CATCGACAAA GACGGCATGA AGCCGTTGGA GATCGAGATC GGTACCGCG AACCGAAGAA
21061 CCCAGCACTA CACATCTCTG AAGAGATCAA GCGCGTCAAC GCGGCTCTTG GCACTGCGGG
21121 GATTCTCTAA ACCGAAAGGC ACGCCGCATG ATTCCCTCAC AAGAGTCTCA CAATCCGAAC
21181 GACCCGCGAC AGCACGTCAT GTGGGCGCTA CGCAATCTCC CGATGATTGC TGGCGTCGGG

21241 GCGATCACGC ATCCGGGTTA CCTGGCGGAT TGGTCAGAGC ACTTGTGGAA GTGCGGCTTT
21301 CGGCACGTCG ACTGGGCTCCG GGAGCTGGCT GATGAGGACG GCAACATCCA CGTCAGTCAG
21361 CTTCCTGACC AGGAGATCAA GTTTCAGCAG CCCTTCCGGG GCCAGCGAAG CGACTACAAC
21421 AACGCAGCTC GATGGGTCCG CAAAGACGAT CCTGACCCAG AGCCCGTGCG TATTCCAGAC
21481 ATTCGCAAGC TCACAGACCA GGAGAACAGA GCGATGATCG CGCAGTACGA ACGAGACGGT
21541 TGGATCAAGG ATGGATCCCT CGGCCAGCG ATAGCCGAGG TCGTGGAGTG ACCCCGTTCA
21601 ACCCAGACTC CATAGGCGAC TACGTGACAC TGCTCGGCGT TCGGTTCTTG ACCTTCTCGG
21661 TTCCCGCATG GTTCAACCGA CGACACGCA AGCACAGCAG TGACATCGGC GAAATCAAGG
21721 AACAGGTATG TAACACCCAC GACACGAACC TGCGCGATGA CCTCGACAGC GTCAAGGCAG
21781 ACATCAGCGA CTTGAAAGAG ATTGTGTTGC AAGGGTTCCA CCAGGTGAAC GAGTCGATCA
21841 ACCTCGAGCG CCGTGAGCGG ATCGAAGGAG ACCGCCGAAA GGAGGTGCGG TGACCTACCC
21901 CACCAACCCA CTAGAGGCCA TCGGCGCTGA CGCGCATTC GAGATCGGTG GGGCGACTG
21961 GAGCTTCGGC CAGGACTACA CCGAACAGGC CATCCGGGCT CTGTTACGA TGCCAGCGGT
22021 CACGATGGAG AACGCTCTCG GCCTGCTCGA AGAGCACCTG CTGAAGCTGC CTCTGGAGGC
22081 GCTGCAGGGC TTCAAAGACA TGATCCCGGA CTGGGTGCGA GGAGCATTCG ACACGGTCAC
22141 CGGCGCTGTG CAGGCGATCA TGAACCGCGT CCAAGACGGC CCGCTGTTCC TGAAGTTCGC
22201 CGAGTTCCAG CTCTTCTGCG AGCGTCTGCT GAACAAACCCG GCCGAGGTCA TCGGCGAGAT
22261 CCCCGAGACG TTGATCGACG GCCTACAGSA CGCGCTCAAC ACCGTCAACA ACACCATCCA
22321 GACCATCGTG GACATGCTCC TGCAGGCGCT GGGCATCAAC CCGGAGGGGG AGCTGATCGA
22381 CCGGATCTTC GACCTGAGCG ATGAGATGGA GTGGCTGCG ACCGCAGCCT CGAATGCAGC
22441 TACCGGCATC CAGGACACCT GGAACAAGTT CTGGGGAGCC CTCACCGGCG GCGTCCCGA
22501 CCAGGACCAG ACCGTCTGCT AGCCCGCCGA GCGTATCGGC GAGCTGGCCG GCACCACGTC
22561 TGCTAACTCG TCTGCTATCG CCGAGCTGCA GCGTCTGCTG GACAACCAAG AGAACGCTGG
22621 CGGCGTGCGC GCGGCTGACG ACTTCGAGCG ACTGAAGATA TCCGGTTGGG ACATCAGGTA
22681 TTCCAACGGA TCCAGCGGCC GAGGGTACTA CCGTGCCGAC GGCACCAAC TGGTCTGGAT
22741 GGACGAAGGC AACGAGCAGA ACACCGCGAC GTTCTCTCCG ACCAACCCCG CAGACGAGAA
22801 GACAGCCACC GACTACCAGA AGATGACGTT GGTCTGTCGG ACTATCTCCG GTGAGGTACA
22861 GACCGTGTTC CCGCCGCGAG GAGGTTGCGA CACCCGGGTA TGGGTCCGCG TCAACGACAA
22921 CGCTCCGACC GTCCGGCATCA CCGACGGCGT GTTCGTAGAG ATCGGCGGCG TATCGAAGGC
22981 CCAGATCGGC TACCGCCGCA ACGGCAATGA CACGTTCTGT GGAATCTATG CTGACTGCAC
23041 CTGGGGTGCT GGATCGATCT TCGCTCTGAC CGCCGGCAGG GCCAACGGTG CTGAGAAGTT
23101 CGAGGTCTCG AAGAACGGCC CCGTGCTGGC CACATGCTCG GACGACGGCG TCGTCTCCCG
23161 GATGGGTGCG AACTACCGCC GCTGGGGCTG GGAAGGCCAG GCTCGTAACC GCAACCTCGG
23221 CCAGGGCACT CCGAACTCGG TCACCCGAGT GACGATCACC GACAACGATC CTACCGGCGC
23281 AGGCGGTGGA GCTGTCAACG TCGGAGGAGA TGTCTAGGT GTACTCCCA TAGAGAACGG
23341 AGGCACCGGA GCTTCGACAG CTTCCGCGAG CCGTACCGCT CTCGGAATCG ATGACCTGGT
23401 CGAAGATATG TCCGACGTAG TTCGTGGATC CGTCGAAGGA CTCCCGTTGA TACCGAAGAT
23461 CTGGGTAGGA ACAGAAGCTC AGTACACGGC TCTCGCCACC AAGGATCAGT CCACGCTATA
23521 TTTCAGGACC GCTTAATGAC TGGTATCTCG ACCGATCCG CAACCTCTCG
23581 ATATTCTTAG CCGTCAGCAA CAAGATATTG AAGGTCAGTC TAGGCACAGA AAAGGTCTGG
23641 CCGTGGTTCA CCGCGGTGCT GACCACGTTG GCCACGGTCG GCACGTACAC CTACAACATC
23701 CCGGACGGGG CCAAGTTCAT CGACGTATC CTCCTCGGAG GAGGCGGGCG GGGTAAAGGC
23761 ATGGCCCTGG CTGACGGCTG GGGCAGAGGT GGAGACGGCG TCACCGGGCT CGTCGGAGCT
23821 CTCGAACGCG GGGTACACAT CCCGTTGTG ACCAAGACGA CCGGAGGCCC TGGAGGAAAC
23881 GGAGGCGCAG CGGGAGCTGG CTCTGTATTC TCAGGCAAGG CCGGCGGTCC CGGAGGCTCT
23941 ACCACGGCGT CCGCTGTGCG ATGGTCAGGT TTGACCGCAA CCGGCGGTCC CGGAGGCTCT
24001 GTGATCGACA TCCTCAGCGT CGCCGGAAAG TCGCCTGGAG ATCGGACCTA CAACGACCAG
24061 CTCTACATAG GCGGCGCACA ACAGAATCA GCTGGCGGGA ACGGCAATGC TCCTGGCGGC
24121 GCGGGGGCTG GTGCCAGGT CTCCGCACAG AGCGGCGGTG CTGGCGCTCG CGGCCAGGCG
24181 TGGTTCTTCG CGTACTGACA AGAAACCCCC CTCTTTAGGA CTCAGTGTCC TTGGGAGGGG
24241 GGCTTTTTTG GTTTCAGGAG GTCTTGGCCA GCTTGGACAT CGCCTCAGCG ATAGCCTCGT
24301 CGCGGGCCTC AGACGCCATC TGGTACTTCA TCGCCATCCT AGGAGTCGTG TGACCGAGAC
24361 GGGCCATCAG CTCCTTGCTG GTCGCACCTG CCTGAGCGGC GAACGTAGCG CCGACAGCGC
24421 GGAGGTCTGT GATGCGGAGT TCCGGCCGAC CGATCTTGCG GTAGCCACGC TTCAGCGACT
24481 TGGTGAACGC GGACTTCGAC AGCCGGTTGC CCTGCGTCTG GTTCACCAGG AATGCCTCGG
24541 GGCCCTTGTT CATCTTCGTA CGGTCTTTCA TGTGCGCTCG GATCATCTCC GCGACGTGAG
24601 GCGGAACCGT CACAGGACGC TTCGACCGGA CGGTCTTGCG GTTGCCAACG ACGATCTTGT
24661 TCCCCACGCG GGAAGCGCCA CGGCGCACCC GGAGCTTCAT CGTCATGCCG TCGTCCACGA
24721 TGTCTTTCG CGGAAGCTCG ATCAGCTCTC CGAACCGGAG GCTCTCCAC GCCAGGATG
24781 ATGCCGCGAT CCGGTAGTGC TCGAAGATCT CAGCGGCGAC GATGTCCAGC TCCTCAGGCG

24841 TCAGCGCCTC TACGTGCGGC TCATCGGCTG CCTTCTGCTC GATCCGGCAC GGGTTCTCTG
24901 CGATCAGCTT GTCCTCGACC GCTGTGTTCA TCACCGCCCG GAGGACGTTG TAGGCATGCC
24961 GGCGGGCAGT CGGGTGCTTC CTACCCATCC CGGCCACCA CGCAGCACC AGAGCTGGCG
25021 TCATCTCTGT GACCGCCACT TCACCTAGCA CCGGGTAGAT GCGGCGCTCC GCGTGCCCGC
25081 TGTACAGATC CCTGGTGCCG TCTGCGAGGT CGCGCTCCAC GAGCCACTTC CGGGTGTACT
25141 CCTCCAGCGT GATGGCGCTG GCGGCTGCGT TCTTCGCCCC GTCTGTGGA GGGGTCCAGG
25201 TCTCCATCTC GATGAGCCGC TTCTCGCCCG CGAGCCAGGC TTCGGCGTCC ATCTTGTGT
25261 CGTAGGTCTG CAGGCGGTAG TACCTCACAC CGTCCTGCGG GTTGACGTAT GAGGCTTGGA
25321 TCCTCCCGCT GGGGTGAGTC TTCAGCGATC CCCATCCGCG ACGTGCCAAC TAGGTCTCCT
25381 CTCGTCGTGA ACAAGGCTAC CGGGTTGCAA CTCCTGTGCA ACTCTCAGGC TTCAACGCGC
25441 TTCTACGACC TGGATTTCT TTCCACTTAG AGGATGCAGC CGAGAGGGGG TAAAAACCTA
25501 TCTGACCCG GCGATATGTG GTCGGCAGAC ACCCATTTCT CCAAACTAGC TACGCGGGTT
25561 CGATTCCCGT GGGGGGCTCC GCTGGTCAGA GGGTGTTTTC GCCTCTGGC CATTTTTCTT
25621 TCCAGGGGTC TGGAACTCTT GTGCGACTCT TCTGACCTGG GCATACGCGG TTGCAACGCA
25681 TCCCTGATCT GGTACTTTTC GATGCTGACA AACGAATAGA GCCCCCGGCC TCGCGAACA
25741 GACGAGGGGC ATTCACACCA GATTGGAGCT GSTGCAGTGA AGAGAATAGA CCGGGACAAG
25801 GTTCACCGG GAGTTGCAGC GGTGCGAACC CTCGCGTCCG GCGCGCTGGC GTTCGCGCTG
25861 TC GTTCACCG CTCTCAGCGA GCTGGCTGGG GCCAACGGGG TGGCCCAAGC AGAGATGGTG
25921 CCCTTGGTGG TCGACGGCCT GACGCTCGTC GCCACGGTCG CCACAGTGGC CCTCAAGCAG
25981 AACAGTTGGT ACCEGTGGTC GCTGCTGATC CTGTCCACCG TCGTATCGGT GGCCGGCAAC
26041 GTGGCACACG CCTACCCCCA CGGCATCATC GCGATGGTGA TCGTGGCAT CCCTCCGCTC
26101 TGGCTACTGG CGTCGACCCA CCTAACCGTG ATGCTGGCGA AGCAGCACTC GGAGCACGCC
26161 GAAGTACCTG TCTCGCGGCC AGAACCGCG CCTCGGGGCC TGGAGCCCGC TCGCGCTTGA
26221 CTGCGCCCGA CCGGGACAGA AATACATAGA GAACCTATGG ATGTAGGAGG CACAAAAAAA
26281 TACCCCCCGA GCGAGCCCGA AGGCCAGCCC AGGGGGCATG GTTCTGCTTC AGTAGACCTT
26341 GCGAGTCCGA CCGGAGTTGA TCATCGCCAT GATGACCCAG ACGGGCAACC ACATTCCGCA
26401 GGTGATGAGC GAAAGCAACA GGTGCATCGC GTGGTTCTGC CTGACAGGCA TGACAGTGGG
26461 CTGCGGCATC GGAGGAGGCG CGACCGGGTA CCGCGAGCCC GCGTACCACT GAGGTGATC
26521 TTGTTGGGGC GGATACTGAT TGGTCATCCC GACAGCCTAC TTGCGGATGG TCGCATCAG
26581 TCCCTCGACC GACTCGCGCT CCACGCGGAT CAGCCGGGGA CCGAGCCGAA CGGCCTTGAG
26641 CCGGCCGTCG GCGATGTAGT TGGGACGGT CTGTTGCTG ACACCGAGGT AGTCAGCGGT
26701 CTCCTGGATG GATGCTCTCG GGGGCATCAG CGCGTCCCTC CGTGCTTCAT CGGTGTCTC
26761 CCGAACCTCG GATCAGGCCA CGATCCTTGC GGCTCTGGAG CTGTTGAGG TTCCTCTGGG
26821 TGACGGTGCT CAACAGACA TCAGCTGGT TGGCTAGCTG GCGGACGTAC CACATCACGT
26881 CTCGAGTTTC CGCCTGGAGG TCGTCTCGGT TCTCTGGGT GATGACACCG TCTTTATCCC
26941 GGAGGATTTT CTTGACCTTG TTGGCGATCT CGCCGGCTTC GCCTACGAGA CCCATCGTCA
27001 CGTAGGAGAG ACCCTCGATG CTGTGCGAGT CGCCTGCACC GGGGTAGATC GCTGTGTCGC
27061 TCGCGGCGAT CTGGTAGATG TCGACGTGCA TCAGATCATC ACCGGGAACA ACTGGCCACC
27121 GGGCATCTGG ATGAACACCG GGCAGCTGGG GGTGTAGTCC GACGAACCCG TGCCGCCCTC
27181 ACAGGCGGAC AGGCTCAGGG TGGCGGCAAG GCCGATGATG GCTGCTGCGA TGGTCTTCTT
27241 CATCTGTTGC TCCAGTAGCT AAGTTCGGAC TCCAGTTCGC GGATACGCTC CTGTAGCCCT
27301 TGGTTTTCCA GGTACGCCTC GCGGAGGTTG GCCTCGGCGC GGTACGCGG CTCGTCTCTC
27361 GACGTGGCCT CATCGATTGC CTCGTGTAGC CGGCGGATCA GATCTGGGAT GGCACCGTGC
27421 AGACCGCATA TGAAGTCGGC GTCTGCCTCG GAGAGGTGGG ACGCCACAG ATCTTTGTCC
27481 TGGGTCTCCT GGTGACCGC CCAGATGACG TGATCCTCTA GCGCGTGGTC GGTCTCGCAG
27541 ATAGAAGGCG GTTCTACCTC CTCTGGCATC CAGTAAGTCT TCTCAGCCCC GGTGGACTTC
27601 GCGCACTGCT GGTAGAGGAT GTCGAAGAAC TCGTGGTCCT GTTCGTGCGC GGTAAATACA
27661 GATCGTCTCTC TTCATCCCAT TCGTCGTAGT AACACGTACA GCGCAGCAGC GTGACGACG
27721 CGCACTCGTA GGTGCGGTAG TCGTAGTCAT CCCAGTCGTC TCGTCCATC TAGCTGTACT
27781 CCTTCATGAT TCGGTGGAAC GCACGCGTCT GCACGCGCAT CTCCAGGTCT ACCGTTCTGCT
27841 TCAACCACGC CCATTGCGCG TCGTGGTTGA TCTCCCACTG GCTCTTGAAT GTCGCTGTCT
27901 CAACGAGGAA CTCGACAGTC AACGTGTGCA GTCCGTTGTT GCTGGGCTGG AATCCGATAC
27961 CGTCCTCAGC GATGTACAG GGCACCTCTT GCGCGTCGAA GTAGACGGCC TTGTCGGTCA
28021 CCACTACTTC AGGGAAGGTG TGCTCGGTC ACGGCGTCCC AGGTATGGGA TGACGCTGGC
28081 CCGGAACCTCA AGGAACACCA TGTGTCCCG GCAGTCTCTG GGGACGTTGT CCGGGCGTTC
28141 GGGGGTGTAG ACGCGATCT CGTTGCCCTC CAGGGTTCCA AGCTCGTTGA GCTTGTAGAT
28201 CGCCAGACCC ATCAGCTCTT CATCGAGACC GTTCGGTGCT GGCAGTACA CTTGGCTTG
28261 TGGCATTAGC CCTCCCTCGG AATTACGTAT GCGCTGAAC CGACGGCCGT AATGCCGTCT
28321 GGCAGTTGGA ATCCGAACCG CTCTCGAAC TCCTCGTTGG TGATGGGGCC GTACTCGAAG
28381 GTTCCGGGCA CTACCTCGCC CTTCCCTCTG ATCAGGAGGT ACGCACCGGC GCGGTACACC

28441	TCCTCGTCGT	TCGGCCATCC	GACTACGGTC	CCGAGGACCG	TGAACTTCCT	CGGCTCCATC
28501	AGGGCACGTC	CACTTCGTTG	ATGAGGAACC	GCATCGGAGG	TGGAGTGAGC	ATTGCCCTCGG
28561	CTATGGCGAT	GAGGGCGTTC	AAGTGACCCCT	TCAGCAGCTT	CTCCTCGTCG	CCTCGGGGAA
28621	GGTGGCGCAC	TCGGCGCTCC	ATCTCCTTGG	CGCGTTCCAG	ATATTGCGGTG	GCTGTCAAGT
28681	TGTCTCTCTT	AGTAATCAGC	GCCGTAGAGC	GAACCCACG	AACGCTTTCC	GACCTCGGGG
28741	TCGGTGCCAA	CCAGCAGCGG	ACCCATCTGT	TCTTGATCA	GGTGGCCAAT	GTGTGCAGCG
28801	GCTCTCTCAG	CCTCTGAGGC	GGGCAGAGAC	GCGACGATCT	CGTCGTGGAT	AGGCAACCGT
28861	AGGTACGGGG	TGTATCCGGC	CTCGTGGAGG	CGAATCAGAG	CCCGACAGGT	CACGTCCCGC
28921	GACGACGACT	GGATCATGTA	GTTTCAGCGC	GAGTATGTCC	GCGAGCTGTC	CACCGGCAGC
28981	CGCCGGCCCA	TCGCGTTGAC	GATGTAGCCG	TTGCGGCCAG	CTTCCATCGC	CAGCTTCTTG
29041	CTCAGCCGCT	CCACACCGGG	GTATGTGCGA	GAGAAGCCCT	CATGAACTCG	CTTGGCCACA
29101	GGGATCGAGA	TCCCCACTGC	CTCAGCGAGA	GCCTTCGCCC	CACCGCCGTA	GACCTTCTGA
29161	AAGTTGGCGG	TCTTCCCAAC	CTTTCGCGGC	ACCTGGGCTG	CGTCAGCGGT	CATCTGGTGG
29221	AGGTCCGCAC	CGTTCTCGAA	TGCCTCGATC	ATGTTGCGGT	CGCCCCACAG	CGCCGCCAGG
29281	ACGCGAAGCT	CCTGCGCCTG	GTAGTCGACT	GAGGCCATCA	CATCGCCTGG	CTCAGCGATG
29341	AAGCATCGCC	GCACGATCCA	GTCCGACGAC	GGCAGCGTCT	CGCGCCGGAT	GCCGGTGATC
29401	GACATCGCG	AGGTCCGCGC	CTGCACTGGG	TTGATGAACG	TGTGGCAGCG	GTCTCTCAGG
29461	TCCCTGGTGT	CGATGAACCT	CTGGACCCAG	GTCTTCCGCC	ACTTCCCCAG	CTTCTTAGCC
29521	TCCTGAGCGA	TGGCGGCAAG	CTCGTTGCCA	TCTTCGACCA	GCTTGTGCGG	CAGAGCCGCG
29581	TTGACCTGGC	GCTTGCCAGT	CTCGGTGCGA	CCGGTGATCT	TGACGCCCCT	CTCCTCAAGC
29641	CCCTCGGCCA	GATCCTCGGT	CGAGTTGACC	TTCTCCACGC	CGTACTCGGT	GAAAGCCATT
29701	GCCTCCGAGA	CCTCCTGATC	GGCCAACCA	TTCTGGGCGA	GCGACCCGGA	GTACTCCACA
29761	TCGAGCAGGA	AGCCCTGCCT	GTGATGTAG	CTGCAGATCT	CACTGATCTT	GTGCTCTAC
29821	GGCACCAGCG	ACCGACTCAC	GTGGGGCACC	AACGGTGTCA	GGCTCTTGCA	GACCTTCGCG
29881	GTGACGATCG	TGTCCATCCG	GGCGTACAGC	AGGTACTCCG	GGTGGAAACG	GTGATGGTTC
29941	GACCAATCTT	TGGCCTTGGT	CGTCTTGTGC	TCGGCGGCTA	GCTTGGCCAT	GAGCTTCTTG
30001	ACGTTCTCGG	CCTGGTCTCT	GGAGATGAAC	TTGCGGATCA	GCTCTTCGAG	CGAGTGCCCG
30061	AACCCGCGCG	CCTCGAAGGG	CCGGGGGTCC	ACCAGCTTCG	CCAGGATCTG	CGTGTCAAGC
30121	ACGCGGGGCC	ACAGACCTCT	CATCTCGATC	CCGAAGCACT	GGTCGAGCAC	CTGGAGGTGG
30181	AAGGAGGCGT	TCYGGAGCAC	CATGCGCTTG	AGAGCGGCGA	TGGCGATCCG	CACGTCTCTG
30241	ATGAACACGT	CTCCCAGCTC	CACCGGCACC	ACCCAGGCTT	CGTCTGAGT	ACCGAATCGG
30301	ACGAGGCGGC	ACTCGAAGGT	GTGCTGTGAG	ATGTCCAGCC	CGGTGGTCTC	AGTGTGACG
30361	GCGAGGCAGT	TCAGGTGAGC	CCGGATGAAG	TTGCGGAAGC	CTTCCAGATC	CTCTGGGGTT
30421	TCAACGACGT	TGACGGTGAC	GAGGTCTCCC	TGAACCTCAT	GCCGCAGCTC	GATCAAAATG
30481	CTCTCTACT	GGAAGTACTG	AGGCAGGAATC	CAGGTGGCTG	AGGCCATCTC	CTTGATGGCC
30541	TGCTGCATGG	CCGCTTCGAA	CGGACAGTCC	GGGTGATGT	CCGGCTTGTA	ATGGGTGACG
30601	ATGATCCGGC	TGTTGCCGCG	GAAGTCTGGG	CTGACCAAGC	CCTTTGGGGG	CCTTCTCTTC
30661	AGCGCCTTGA	TCAGTTCCTC	AACCGTGGTC	CCGGTAGGGG	CCTTGCCGTC	AGGCAATGCC
30721	TCCCTCTCGT	ACGGCAGGTC	CAATGGGATC	GTGTACCGCT	CAACGTCTTT	GATCTTCATC
30781	GAGCCTCTTC	CTCTTCGACT	ACCTCGTCTA	CCCGGCGGAA	TAACCTCCGT	AGTTCTGCGG
30841	GTAGCAATAC	TGGGTACTTC	TCTCGGGCTT	CCTGCATCGC	TACCGGATC	CCAATCAGGG
30901	CAGCGAGCAG	TTCAATTGACG	GAGTACGCCA	ACAGCTCTTC	GCGGATCTCT	TCTCGGGTCA
30961	TTAGTGGTAG	ATCCCCCGGA	CGGTGCGCGA	GATCGTGGCA	GGGTTCACGC	CGTAGTTCTC
31021	GGCGAGATCC	TTCTGCTTCA	TACCGCCGAG	GTAAGCCTGG	CGGATGTCTT	TGACCTCGCG
31081	CTCGGTGAGC	TTCTTGCCTG	TCGGCCGGCT	CGGCCGGTTC	TCAGGCTTGA	CCTGAGCCAG
31141	CGCCTTGCCG	AACAGCTCGT	TCTGCTGTCG	CTGCTTGATC	GCGTACCGAC	GGTTCGCTGC
31201	AAGCACCTCG	TTGAGCCGCT	GGGACAACTT	GACATTGGCC	TCACGCACTA	CCTCGACCTC
31261	TCCGAGCAAG	TTGCTGATCC	GGTAGTCCTT	GTCTTGGTTC	TCGATGGCCA	ACCGGTTGTT
31321	CTCCTCGGAA	AGCATCGAGA	CCTTGTATTG	CGCCTCTCCC	AGCGCAGCTT	TCAGGTGCTT
31381	CTTCTCTATT	CAGCGCCCTT	CTCTCGGCGG	AACTGTTCTG	ACTCGTCTTC	GGTCATGTAG
31441	TAGTAGTAGT	CAACGACCTT	GTCCCAGTTG	AAGGTTCCGG	ACGTGCCGTC	ATCGAACCGG
31501	ATGATCAGGA	CACCTCTTGG	GGTGTCTAGG	ATCGGCTCGC	CAGCCACGAC	GTGGAAGCGG
31561	TCCTCGAGGG	TCACCGCAGT	CGCTCTGCGT	GCCATGTGAG	TTCTCTCAG	TAGCTGTAGG
31621	GGACATCCGG	GATGTCCTGG	TAGGTGTTGG	GTGCGATCTG	TCGGAGCTGC	CGAAGCAATT
31681	CCCCTGCCAG	CTCACGGATC	TCGGCATCCG	CGGCCTCGTG	CCAGCGGGCC	TTGATGACGT
31741	ACCGCCACGC	CCGATGGTTG	CCCCTGACGA	CCATCGGTGA	GTTCTGTCATG	TTCCGCAAGGA
31801	CAGCTCGCGC	TGCCTCGCGG	GCCTGCTTGC	GCGGCAAGCC	CCGGTCAGCC	AGCCGGTTGA
31861	CGATGTGTTT	GTAGACAGCG	TCAATCTCAG	AGCTGACGGA	CTCCATGATG	TGGACGAGGT
31921	CGTCTCGGTC	GTGCGGGTGG	AGCTTGAAACA	GAGCCGGGGG	CAGATGGATG	CCAAGGTCGG
31981	TCGGATCCAC	ATATCGCTGA	GACACCACCG	AGAAGCTCAA	GTGACGGTGA	CGCTCCAGCT

32041 CGGTACAGCAC CGACCTGCTG GCCTCGATGT AGAACGTCCG CGAGGCGTGC TCGAACACGC
32101 TCTCGTGGCC CAGATCGATG ATGTGGTTGA GGTAGTCTC GTTCTCGGCA GTTGCCGGGT
32161 TCGGTCGGTG GAACGACCGG TAGCAGTTCC GGCCCGCGAA CTCGGCCAGC TCGTCGGCAT
32221 CGAAGTCGCC GAAGTAGGGA TCTTCGTCTT TGGATTCTTC GAAGTCATCG ACCTCGAATC
32281 CGATGTCCCG CAACGCACCC GGATCGATCT CGGTGGCAGC GATCAGTTTG GCTTTCATAC
32341 TCTCCGCTCA GAGTTGGTGG AACGAGGTCA GCCAGGGGGC AGCGAAGCCC TTCTACAGCT
32401 CCGCTTGGCT CGTTACCGGC TTCTCGACCT CGGTGGATGT CAAGTAGTCG AGATGACTAC
32461 TTCTGTCCG GCCATTGCGC GTCACACTGC TGATCGCGAG GTGCGGTGCA GGAGAACAGC
32521 GCGTACGGCT TGCCCGTCTT CTTCGAGACG CCGGACTTGT AGACCATCTC GCCGTGCTGG
32581 CAGTACCGCT TCTGCCACC AGGCGCTTCC TGAGCTGCCT GCGGGGCGCG AGACTGCTGC
32641 TGGCCACGGC GCGCGCCCTT GGCCGGCGCG GATCCACCGG AGCCTGCGTA GTGGCCTGCG
32701 ATCTGCTGGA CTTGTCCAT CAGCGCTTG AACTCGGCG TGTGACCTT GGCCAGCAGC
32761 TCGGCGGGT CCGCACCTT CACGACACC CACGGTCCG TGTACTGACC GGCGAACTTG
32821 AACGTGGCGG ACACGCTATC GGTGGAGTGC TGGACCGCCA TCGAGTCGCG CACAGACGCC
32881 GAGCGCTCG TCACCTGCGC CGACGGCGCG GTCTCAGGCT CAGGAGCCGG GGCCGGCTCG
32941 GCGTGGGCG GCGCGGTGCT CCACGGATCG TCGTAGGACA ACTGGTTACC TTTCACTTAA
33001 TGGGCGATGC GCGGTGCGG CACTCTTCAT CGACACCGTC TTCGACGGCT TTGGCCGCGC
33061 CAGATCGTA CTGCTGCTTG GTGATTCGCT CGTACGGAGC CTGCGGGAAG CTGGACTCCG
33121 GUAAGATCGT GGAGCGCTTG ATGAGCGCCG CGAACCTCTT GAGATCGGCT GCGACATCCT
33181 CGGCTGCTG GCGGCTGGA TGGAGTTCG CGGTGAACGA CACCGCGTTC TCAGCCAGC
33241 ACATCTGCTA GAGCGCTTG AACGCCAGGA GCTGGTGGAG GGTCAACTCG TCGGCTGACT
33301 CAACGATCTC CTCTCCCA CCGAGTTCTT CGACAGCCTG GACCAACGTG TCCTTGGTGC
33361 GGATCGAAGC CACCTCGCTG TTCTGCGGGA AGAGATCCTT CTCGATCTCG TAACCCTCGG
33421 CTGCCAACCT CCGCAGCTCG GCGATGTCG TGTGAGGT GAACCCGACA CGCCGGATGA
33481 AGTACCGCGA GAAGATCGGC TGGATCCCTT CGGAGACTCC TGGCATCTTC GCCACCGTGC
33541 CTGTGGGAGC GATGTTCCG TTCTTCACCG GGACAGGGAT CCTCAGATCA TGGCGGAACC
33601 GTTCGGCTTC TGAGTCGACC TCAGCGGCGA TCTCCGCAA GAACTGGGTG AACCCTTAT
33661 CTCCGGGTGC CTCGGAGTAC CTGTTACTTG TGAGGGGCAA ATAGGAGGCA ACTCCGAGAT
33721 GACCCACGCC SATSCGACGG TTTCCGTCCA GAACCTCCCG GCTCTTCGGG TCGGCCACTT
33781 CCGAGAACGT CGCCCGGATC AGGAATCTCG TCATCAGACG ATGCGCGCGG ATCAGGTCCG
33841 GGTAGTCGGT CTTGCCGCGC GCGCTCAGCA ACGCCGCGAG GTTGATGTGG CCGAGGTTGC
33901 ACGGCTCCCA CGGTTGAGA GTGATCTCGC CGCATGGGTT GGTGCAGACC ACCCGGTTGG
33961 GCTCACCGAC GTTCGACAGT GACGAGTCCC ACATCCCGG CTCTCCGTTG CGTACGGCTC
34021 CCTCGGAGC TCCCTGAGC ACTCGGTGGG CTCGCTTCTG CTTGGGATCG TCCTCGCGGG
34081 CGACCGCGAA GCTGCGTAG CCTCTCTTGG CCAGACGCGA GAACTCGTCG TCAACCTCGA
34141 CCGAGATGTT CGTCCTCCAG TGCTCGCCCG TGCTCGCCTT GATGTTGATG AACTTGTGCA
34201 TCTGGTAGTC GTCCGAGTGC ATCATCGACA TCCGCGCGCA CCGGCGCACA CCGCCGCGCA
34261 CAACACACTG AGCGATGGCG TGGTCGACCT CCATCGCGGC GATGCGGTGC AGCGTGATCC
34321 CTGCGTACTC CGAGAAGATG TTGGCGACCT TCTGCGAGAT CACAGCGAAC GCGAGCGGGC
34381 CGCTGGCCAC TCCACCGAAC GTCTTGAGCT TGGCCCTTGG CCGCCGGATG CCGCTCACGT
34441 CGTACACCGG CTGGTAGTGG ACCGTGCCGG GTCGGTAGTG CGTGTGATC AGATCGACCA
34501 GCGCAGCAGC CCAGCCCTCT CGTAGTCTT CGATGGCGTA GGCACCGGCC CAGTCGTGGC
34561 TGTAJTGCTC CGACAGAATG CCTACATCTT TCATCGCCTG GTAGTGGACA TGCTCTGGAT
34621 CACAGACGAT CTCGACCCGC AGGGGGTTTA CGACCTCGGG GTAGCCTTCG AGGTAGTGGT
34681 TCGAGTAGTT CGCCCCGACT CCCCCGCCCT CCATCAGGCG CATGAACGTG AACTGGAAGT
34741 GGTCCGAGAT CTTCTCGGGC CAGCCAGCTA CCCAGCAGT GAAGAGGTGC TCGCGTTCT
34801 TGACCCCGGA GGCCACAGA TGCCGACCTG CCGGCAGCAC CTTGAACCTG GTCATCAGAC
34861 GAACGAGATC TTCTCGCTCT CCTTCCAACA TATGTGCGCG GTCGACAAGA GCAAGATTGC
34921 CGTCCACGAC CCTCTCGACC GTTTCGGGCC AGGTTTCTCT CGAGCCGTCA GGCTTGGTCC
34981 TGGCGTAGGT TCGGTTGTAA ACGAGTTCAC CGGTTGGTCC CCAAGGGATT TCGTCAGTCA
35041 ACTACTTCTT CTCAGTCAGT TCGTATCGCT TGAAATAGGC GTCGGCAGAG TCGCCGCCAG
35101 AGAACGAGAC CCGCTACTCG ACCGGGCTTG CACCACGCAC CTCGCGAGTA AGCAGCCCT
35161 TCCTTCCCGG GAACATCGGC CAGGTTCCCT TGGAGGGGTG CTTGGTCTCG TCCCGTGGGA
35221 CGATGACCTT GGTGCCCTTC TTCATGCCGA CTTCCGTTCT CCGTAGCCGG GAGTGAAGCA
35281 ACCCCCGACG TACAGCTCGA GATCTTCTTG CGACAGTTC TCCAGTCCGA TCGGCGGCTG
35341 GTGCGGGAAC AGCTCCGGGA ACACCTCGGC CCGGTACAGC TCCGAACCGG GCATCCCGTT
35401 GAACGTCCGA TCAAGAAATG TGTGATCGGC ACCTCCCTCC CAAGAACTCG GAGATCGGCG
35461 GCTCGTAGAG GTAGCCATCG CGCAGCTCGG GGTTCGATG GAGCATGATC GCGTGTTCG
35521 CTGTGGGGTC AGAGTGCCCA TCCCCCTGCG ACTTTCGGAT GTCTGGGAAG ATAGCGTGCT
35581 TGCTGCCCCG ACCATCCTTG ACGATGACCT TGCCCTTGTC GTCCTTCTCC ACGCCAGCCG

SUBSTITUTE SHEET

35641 TGATCGCGAT GATGTTGACG TGCTCGGTCA GCGACTTGTG AGCGCGGAAC AACCGGTTCT
35701 GCCCCGCTCTT ATCCTTCGGG GAGATCCCGT CGGTGTAGCG GCTCCTGATC GCCTCTGCAT
35761 AGCCCCCGTT CTGAGCGTCC AGAGCCTTCA TCGCCAGCGG GAGGATGTCT ACCAGGTACC
35821 GATTGGTCGA CTCCTCTGCG AGAGCCTTCT TGACGTTCTC GGACGAGTAG TGCTTGGCCT
35881 CCTGGAACAA GTCCGCGGGCC TTGGCCGCTC CCGACAGGAT GTTGCGAACC TGATTGCGTA
35941 CGTAGTGAAC TGCCTCACCA CGGTGCAAGC TCTCCAGCGT CTTCTGGATG TACGGGCTCT
36001 CGAGGTACCA GACCCACAGC TCTTGGATGA TCTCCTCGGC TGTACGGTTG GTCTCCCAAC
36061 CGATCAGCGC CTTCCGGGTG GCCCTGTGA ACAGCTTGCT GATGTCGTCT GTCAAGGCAT
36121 CACCTTTTCGT AGGTACTCCT CCCGGTCCAA TCGGCGGTCT AGGTGTCGAG TGACCTCCTC
36181 CGCGAAGACC TCGCGGACTT CGCTGGAGGT GATCTGGCGC GAACGTGCGT TCTTGTGCAG
36241 GTACGGCAGC TTGGTGGCTG TCAAGTTCTA GACCTCCAG ACTCGGCCGT CGACCGAGAA
36301 CCGGCCCTCG ACAATCGGAA CAAGCTCAGG CTTGACGTGC TGGCCGTCTGA CCGTCAGCAG
36361 AGCAAAACCA CTCTGCCAGT TGGCTGTTGC ACCCTTGAGG TACTGAGCTA GCTTCATGTT
36421 CATCAGGTTG CCGACCTCCA TCGACCACAG CACCTTCTTG TTGCCGCGT AGCCCCAGCT
36481 GTGTGGCTTG ATGCCCTGGC GGTGGGTGTG TCCGATGATC ACCGACGTGC CGAACCCGAT
36541 CATCGCGTTG TACGCGGTGT CAGCGGACTT CTGCGTCACC CGGACCCAC CACGGTGGCC
36601 GTGGGTGGAG ATCCAGCCTG GAGCGATCTT GTAGAATCA GGCAGCACGT CAACACCGAA
36661 CCCGTCAAG TCCAGCAGGT TCTGGAAGT GAACGAGCTG ACGTACTCGA CCAGCGCCCG
36721 GCGGAAGTGG TGCAGGTAGT CGACTGGCCG CGCGTCTGG TTGCCCTCGT GGACACCAAC
36781 CCGGCCGTCTG TAGACCTGGC GCAGCGGCTC CAGGAACCGC CGCTTGCACT GCTCGGAGTC
36841 GGGCTTGATG CGCTGAGCGA ACTCTTCCTT GGTGCCCTTG GTCCACCGAG ACGGGCTCGG
36901 GTAGTCCATG AGGTACCCGA TGTGGACGAC CTCGTCAAGC TGGGTGTCCC CGATGTAGCC
36961 GATGACCGCC TTCAACTGCT TGCATCATC GAACGGAATC TGGGTGTCCG AGATGACGAC
37021 GATGCGCTTG CTCACTCAGC GACCTCGGTG AAGGGGCCCC GCATACGTTG CTCGTGGGAG
37081 CTGGCGTTGC CTCCTGACCA GCGTCTGTTG CCGACCTTGG TGTGGTGCAA CCCGTTGGGG
37141 TAGTAGATCC ACTTCACTCC TGTGGCGTTG GTGACGGTCT TCACATCGGC AGGAACGTCC
37201 AGCAAGGTGT CCCACTGGCG AGGCCCTTG GGATACCGCT CGTCTCGGG GAGCTGCATC
37261 TTCTCCAGAA CGCCTGCGTA ACCGGCGATG TCGACCACCG TGTCTCTGGT GTAGCCGTTT
37321 TCCATGAACC GGGCGATCTT CAGCAGGATC ATCATGACGG CCACGTCTCT CCGGGTGAAC
37381 TCGACCTCCG GCTTGTACGC GCCCCACAGG GTCGEGATGC GTTCTGTGGT CTCCTTGGCG
37441 TCCCCCTAGT CCTGGGCTCG CTGTCCGTTG ATGATCTCTT CGGCGGTGGT CAGAATGCTC
37501 ACAGTCCAGT CTCCGATGCG GTGTAGTAGT CGATCAGTCT ATCGAGCTGG TCCGTTGAT
37561 AGCCGAGGAT CGGCTTGTGG GTGTCACTGA CGACGACGGG AACCGACATC GCGTTGAGCA
37621 CCTTGGTGAC GTAGTCGTAC GCCTCCGAGT TGGCCGTGAC ATCGACTGCG TCGAAGTCGA
37681 TCCCGGCAGC CGTCAGCTTG TCTTTGACTC GCTCGCATGG CTTGCAAGCC GGACGGGTGT
37741 ACACCGTGAC CGGCGCGAAC AGCGTTCTCA CGTGAGCACC ATCCAGTCCG ATGTATCGGT
37801 CTCCATACAT CAGATCCTTT CCAGCAGAGC AGCTTTGCCC TGGCATGTGA TGGCTGAGTT
37861 GACATCCTCG CTTTCTGGCA TCGGGATGAT TCGGGCGTTC GGCAGCGTCT TCGCCACCGA
37921 CCGGGCGAAC TCCATACCGG CGTCGTGCGC GTCGGCCAGG ATGTTACGCT TGCGGTAGCC
37981 CAGGAACAGC TCTCGGAAGT ACGGCTTCCA CTTCTGGGCT CCGCTGAGCC CCACCGTCCG
38041 GAGCCACAC AGCTCGGCGG TGTCTGTGTC GAGTTCTCCC TCGCATGCTC CATGTCTCTT
38101 GCTGTATTG GTCAGCGCGT AGGTGTTGTA GAGCCGGTCC TTCTCCCTTG GCATCGACAG
38161 GTACTTCGGT GTGCCACCGT CGATTGCGCG ATACCGGATC GCAGCTACCG TCCAGTGACG
38221 CCAGGGCGAC CACCGCATAT ACGGAATCGC CAGGCAGCCC CGGTACATCT CATGTCCAGG
38281 GAGTGGGTCTG TCCACGAATC CCAGACCGAA CCGGCTTAGT TCCGCTCGC CGGCCAGCCC
38341 GCGACTCGCC ATATACTCGT CGGCTGGGCT TCCGGGCAGG CTTTCTCTGT ACCGGGACGT
38401 TGCCCTCCAC AGTAGGTTG TCTGCGATTG GCTTAGCCTC TGCAATGTC ACCTCCTCTT
38461 CGTGACGAAT GATCGAGATC ACGTCTCCAC GGACCCGCA GGCCATGAG TTGTAGCCCT
38521 GTAGGTCGTA ACTGACTGCG GCAGACGGCG TTTCTGCGC GTGGAAGGGG CACAGGCACT
38581 TGTTCACCTC GTGGTGGTCA GGTGGTGGTT CCCAATCCGG GTGGTAGCGA AGAATCGCCC
38641 TCGCGATGGG CGAGTCGTTT ATTCTCTCTC GTCAAGCTCC TCGGGAGAGA GCCCTTCGAA
38701 GATCCCGTTC AGGACGGCGG CGAAGCCCTC GCCGGTCTCC GCTGCGTCTG GCATCTCTGC
38761 AATCGTCTTT GCCATGTTTC CTCCTGGTGG ATGTCAAGTT CGAGACAGCT TGTACGCCCT
38821 GACTGGAGCG ATGCGCTCCC CGATGACTTG GACGGCCGGC GGGTTTCAGCA GGTACTCGAT
38881 GGCCCGTTTG AAGAACTCGA TGCAGTCCCT CGCCAGCCC AGCGTGTACT TGTGACAT
38941 CGTGACAGAG AACCCTCGGA CGATGCCTGT CTTGTGATCG TGGTCGACCG ACAGGCGCTT
39001 CTTCTTACCG TTGGCTCGCT GGCAGATGTA GCACCGACCA CCTTGGAACT CGTAGACTG
39061 CCAATACTCA TCGCCGGTGA TGCCGTAGTG GGCCAGGATC CGGCTCTCCC AGCTCGTAGA
39121 GCTGCGAGCC GTCCTGAAC CTGCGGTGATG AGTAGCGCAT CGTGCCCTG GATACTTGGC
39181 GTCTCGCGTG AGCGGGAGCC CCTGTGCGAC ACAGTCTTTG CAAGGCTTCC GCTTGTGCTT

39241 ACCGTTCTGC ACCCGGTACC CCGGAGACCT CTTCGCCGCC CTCGGCACGC GCGTCTCCT
39301 CCCGGTTCTC CATCACCATG CAGAACCACG ACAGCAGCCC TGCCAGGGAG ATGTAGAAGG
39361 CCACCAGAAC TTGGCCGCTC ACTTCACCAT TCCTCGAACC CACCAGCGAG ACAGCGCCTT
39421 ACGCCCTTTG TCGAGCGGGG TCAGCTCGCG CTCATCGTCC TCACCGAAGT CGAACTCGAT
39481 GCTGGCGATC TCGTAGCCGA GGATCTTGAA CGACACGTTT ATAGGCGGTC TCCGAAGTTG
39541 ATGACGGGAA TGCCGGCCCT TTCGGCTCTT CGCATGCAGT GCCGGGTGCC GACTGAGTTG
39601 CCGAGGGGGA ACGCCAGACA GATGTCCGCA CCGGCCCTGA CCATCTCGAT GTTGCGGAGG
39661 ATGCCAGCCC GCTTGCCGTA GCGTTCGCGT TCGGCTCGGT GCAGCTCGGG GAGCACGTCC
39721 CATCCCTCCT GCTTCATCCC CCAGGCCAG CCGTCTGCGA TGTGCTCAGC GCCCGAGCG
39781 CCGCCGTGGA CGACCGTGAG ACCGGAGAA GACCGGTGGT ACTCAGTGGC CAACGCTTCC
39841 CAGACCGTGG TGCGGTCTT CCAGATCCGA GATCCGGTGA TCAGTACTCG CCGCATCAGA
39901 TCGCCTCCA CTGCAGGCG TCGTCCGAGC TGACCAGCTC CGCTTCGTAG ACGCCGTAGC
39961 GGGTGGCCAG GAACTGGATC ATCTGCGCCT GCTTGTACCC GAAGGGACAT TCGTGGACGC
40021 CGCTGATCGG GTATCTGACT CCGTATTTCA CTTGATCCAC CGCTTCGCGA TTCGGTTCGAC
40081 GTTCTCCTCG GAGACGTTG GGGCGAGGCC GGTGAACCTC TGGCCGTGGA CCTTGGTCTC
40141 GATCACGCGA GGCTTGCGGG GATCCGGGCT CTCCGGGTG ATCCGCTTGT GGGTCCAGAC
40201 GGTCCGGCTT GTCTTGATCA GAGCGCCAG CACCTGCTGG CGCAGTGGT TGGTCTTGGC
40261 GGGCATAGCG TTTGGAGTGG TCATCTGGAT CCTTTCTCG GTGGCTGTCA AGTCGGTGTG
40321 CGTAGTGAAG CCCCCCAGG CATGCGCGCC CCGCCTGGGG AGAGTTGATC AGCGCAGTTC
40381 GATGTCGGGC AGGATCGCCT GCGGCTTGA GTTGACCTGG TAGAAGTCGG TCGAGACGTT
40441 TGCGCCATCG ACCTGCTCCA TGAAGTAGGA GACGTTGTCC GACAGGCCCA GGAAGTGCTT
40501 CTTGATCCCG TCCTTGGTCT TGCAGGTCA GTCGAGCTTC TTCGACGCGT TGTCCGCTT
40561 GATTGAGCAC CGGCCCTGGA TCTCGAGCAG GTACTTGTCC GTGATCCCGT TGAAGAACAC
40621 GATCCGCGCA TTGATCTCGA AGTTGTACG GGCCTTGTG ACSTTCTCCG ATGCGACGTC
40681 GCGCTCGGAG GTACACGCGG AGAGGCCAG GATCGCCGAT CCGGCGATGA GTGCGGTGGC
40741 GATGATCTTC TTGATGTTG CTACTTTCTG TTTGGTGGAT GTCAAGTTAG TGACCGAGCT
40801 CGTTGATCTG CATAGTGTCT CCGACGAACT CCAAGGAAGC GAAGTCTTGT CCCGACGGGT
40861 CCGACTTCCC CCTTCGGTTC TTGACCGTGG AGACGTTGAG CATGTCCGGG CCGAACCCGT
40921 CCGATACTCG GTGGAGAGTG AGGATCATCT CAGGAACACG CCCGATCTGA CCTTTGATGC
40981 CCGACAACGG GATCGGCTTG TCGCCGCTGT TGTGCGGGCC GGTGACGTGG TGGAGCCCGA
41041 CGACGCATGA GCCTGTCTCA CCGCCCATCT CGTGTAGGTA GTCCATCAGC GACTCCAGAC
41101 CCGAGAACGG GTCGTCTCCC TCGCTTGAAT CCGTGCAGGAC GTTGGTGTAT TTGTCCACGA
41161 CGATCAACGC TGGGAAGTCC TCGTACGCG CGTCATACGC GGCCAGAGCT TTCTCGATCT
41221 CGTCCAACGA CCGTGATGCC TTGTATGTA ACCGGATCGG GATCTCGTCT AGTGAGTCAG
41281 CTACCGCCTC CTCGATGTTT TGCTCGCGAA CAGCCCGCGT AGCTCGTTTC AGCGACCATC
41341 CGCTGAGGAT GGACACCGAA CGGGAGAGCT GGGTGAACGC ATCAGAGTCG GCCGAGAAGT
41401 ACAACGTCGG CACCTTCGAC TTGAGCGCGT AGGCGAGGAC GAACGCCGAC TTCCCGGTGC
41461 CCGGGCCGCG GCAGACCAGG ACTAGCTGGC CTCGTGCGAG ATGTGTACCT TTCTGGTCAA
41521 GCGCGGCCCA GACCGGGGTT AGCGGATCCC CCGCCGACCC TCGGATGTAT AGCGATTGTC
41581 TAGGTGTGTA CACCTTCCTC CTCGTGGATG TGATTGACCA GGTGATAGAT CTCGTGCGGA
41641 GAGACAGCCC GGGCCAGGCG GTCGATCCCC ACGTGGATCT GTCTCCGGTG GATGTGTCGG
41701 GACAGGATCA TCGGCGAATG CGTGTGCCCC TGGATCAGGA TCTTGCCATC GTCACGGAGC
41761 CTCCACTGGG TGTGTGCGTC CTCGCTGTGT TGGTCCCGA CGTATGGGAA GTGGCTCAGC
41821 AGAACATCTG TGTGCCCGCC AGCGTCCCCG TACAGCGGCA CCGGATACG AGCTGCCGTC
41881 GACACATGCT CGAACACCAT CCAGTACGCA CCAACCAGCT TGTGAGCATC GCGGTTTCATC
41941 GGGTGGGGCC CATCGTGGTT GCCCAGGATC AGCCGTTTGC GGCCTGGCCG ATCCGAGATC
42001 CACCCGAGGG CATGTATCTG CCGCTTGGTG GAGCCAGAGG AGATGTCAAC TAGGATCCAG
42061 ACCGTGTCGT CCTTGCCGAC GACCGATCC CACGCCTTCG CCAGGGTGGC GTCGTGCTCT
42121 TCGACATCAT CCGCCAGGTT GCGGATCTCC ATCAGCCGCT TGTGTCCGAT GTGTAGATCG

42181 GACGTGAACC AGGTGTTGCT CATGGCTTCC TTTGAGAAGC GCGGGCCGTA CAGCTCGATC
42241 ACCAGCGCGT GCAGCTCCTC TGCCGCGTCG TCACGCTCGA ATCCGACGCA GGAATCGTGC
42301 CCGTCGAGGA TTGCGACGAT CTGGTCGTAG AGGCTGGGCC TCACTTCACC TTCTTCGGAT
42361 CGATCAAGGC GTCGTGAATC GGCCGACCGC CGCGAGCCGC GTGCGTCTCG GCGTCCAAGG
42421 CTCGCTGCAT CTGGTTCATC AGCCGGGTGC CGCGCAGCTT GAGGATCTTC ATGCTGCCCC
42481 GACCCCTGTA TCCAGCGCGG TGCATCCGTA GGACGCGAGC TGTCTCGTGC GGGGCTATAG
42541 GTGACCTCAG CGACGGGTGG TTTGGATCCC AGTTCGTCTAT GTCTTCCTCT CGGTGGCTGT
42601 CAAGTTGCTC ACAGACCGAA CTCTTCTGTT TACTGCGGGA TGAAGTGGCC GGCCGTTTATC
42661 GTTCGGCTCG ATACCTCTCG CGTACCGAAC TCCTGCCCGT TCCATCTCCG ACGCTCCTCG
42721 AACTCGATCA CGATCTCTCG TCCGGGATGA CGCACGGCCT CCGCTTGGGC AAACCTCGGT
42781 GCAGCCTCTG GGGTCCGGAA CCGAACTTC TCGAGGCGCT ACAGCTCCTG GTGCCACTTC

42841	GGCTTGTCAG	GAATCGGCC	CATTTCACG	TACGTGTAAC	CCGCGTCGGG	GTCGAGTTCCG
42901	AGCGTTTTCT	TGTATTCCTT	CGTGCCCTGCC	TTAGAGGGAA	GGTGAGTATC	GGTGGCTGTC
42961	AAGGTGACCT	CACTTAAAAA	CAGGGCAGCT	GTAATTCACA	TCACAGAAGC	CGCATTGTGC
43021	AGGTTCAAGC	AGAGGCTCGA	AGTACCAGC	CTGGATCCGA	GCCTCGACCT	CATGGAACCT
43081	CTCGGTGATC	CGCTCCCGCG	TCCAATCGGT	CAGGTCGTAG	GGCGCAGTGG	GCTTCGCCTT
43141	GATGCCCTTC	TTCCCGGCA	TGAAGTAGTC	GCCCGTCTTC	GGAGCCTCCA	CGTCATAGGT
43201	CATCGCGACC	GCGAGCGCGT	ACACGCGAG	CTGGAAGTCG	TCACCCGGCG	AGTTGCCGGT
43261	CTTGTAAGTC	CGGACTCGAA	GCTCACCGTT	GACCACGACG	ACCGCGTCGA	TGAACCTCG
43321	GACGCGGATG	CCGTCCAGCT	CGATGTTGAA	CGGAAGCTCG	ATGGCCGGCT	TGGGCTGTTC
43381	ACACTCCTTG	CAGTTGGTGT	CTTTCCACGC	CTCCGTAGAG	CAGATCCCTC	GCCCAGGGGT
43441	AGTCCAGATC	TGCTGGCCCT	TGTCCTCCG	CCACGCGATG	AACTTCTCTA	CGCTCCAG
43501	TCCAAGGTGG	AACCGGCGCT	CGATGTCACG	CTCACCGTTG	TACGGCCCGG	ACCAAAACCA
43561	CCACTCGAAG	TTCCGGGTTT	CGTCGCACAG	TGCTCCGATG	TCCTTGGCGT	ACTCCTCGCG
43621	GAAGATCTCT	TGTGCCCGT	CGAGGCTCAT	CTCGCGGCCC	TCGGCCAGAG	CCTTCTCGTA
43681	GACCTCAGCG	ACGGTGTGAA	ACGCGGTGCC	CTGCGGCAAC	CACGCGCGAG	GACGAGCCCA
43741	TACCTTGTCG	ATGCGAGCCA	GCTTGTACGC	CTGCGGGCAA	CGTGTGTATT	GTTCAACTG
43801	GCTGACGCTT	CGCAGCGGCA	GCAATGTCTT	GGTGTCTGTC	ACGCAGCGGC	CATCCTTCCC
43861	TTGCCCTATCG	TCTCGTTCAG	CGCCCCGTGC	ACAGCGACAC	TGAGCAGTTT	TGCGACCTCC
43921	GACATGTCAA	TGGATCCTT	GGGGAATTGG	TCAGCCTGAG	TCATCTCTAG	CACCATCCAC
43981	TCGGTGCCCT	TGTCGCACTG	GATCAAAAGC	GGATCAAAGC	GAGTTCCCCG	TGCTACGTAC
44041	TCGACTTTGT	TCGCGGAAAG	AATCAAATTC	GACACAGGCC	GATAAAGTCG	TGAGGTGTCT
44101	TTTACACGAG	GACTGCGGTA	GACGAGGAGA	ACTGAGACTG	GGTCTTCGTC	CAGTTGGCCC
44161	TTCCACCACG	CCTCACACCT	CTGCGGGAAC	AGCCACCCTG	GATGATCGGC	GATGACTTGC
44221	GGTGAGGTGT	GGACGAGGTT	GTCTGCGAAT	AGCTTTGCGA	GCCGAGTGAG	GGGCACGGGG
44281	TTTCCTTTTC	TTGCGCGGCC	TGGGTGCGGT	GACACAACCG	GTCGTGACTT	TTAGGGCTCC
44341	GAGAGAAGCT	CCTCGATGTC	GTCTGGCCAC	GACCAGAGGA	GTTCAACCTC	GGCGGTGAGG
44401	TTGGTGTGCT	CGTTCACCCG	GATCAGGAGA	TCGTATCTCT	CGATGCCTCG	GGGACGCTAC
44461	CTGAACCCGC	CGCCGGCCAT	ACCTTTCGTG	GGCTCGATGG	ATGGGTGCAA	CTCGAGCACT
44521	AAGTCGTCGT	CGCGGAGCAT	CTTCAACGAC	GACAATAGGC	GCTTCTTCTT	GTCTTCGGAC
44581	ATCGTGCGGA	AGCTACCCAC	TCGCATGTAC	TCGCCGTGAT	CCCGGAGCCT	CTGAAAAGCT
44641	TTGCACTTAT	CGTGAGGTTT	CCGCGTGTCC	CACGGCCAGT	TCTGCTGGAC	GATCTGCCCTG
44701	GTGGTCAACC	GTCTCCGTA	GGTCTGCTTG	TGCCAEGACA	CCGCTTGTGC	AGTCACGCCA
44761	TACAGCTCTG	CGATTTCCGT	CTGATTAAC	CCCTTCTGTC	GAAGATCTTC	GATCTCGCTG
44821	AGAGTGAGTG	GTATTCGGCT	AGGGGCCGGA	ACCACTGCTT	TGTGTTGGAT	TTTGCCGCTC
44881	ATGTTTCCCT	CCATGAGAAA	GGTGCGTGCG	TCTCCGCCGA	TTACGGAGAC	ATGTTGGTGC
44941	CTGTCAAGGA	TACCCCTAAT	TGATGTTGCT	CTGCGGAACC	ATATTCAGTT	GTGTTCCCGG
45001	ACGCCGTGGC	CGTCTCCAC	TGGGCGTGGG	ATCGACTGGC	GTTACGCGGT	CGTAAATGTA
45061	GCGGCCTGCC	CCACTCGGTA	GCAAACCTTG	TGACAGGTAT	CACCTAGGTC	GCCTTCTGTT
45121	ACACGTTGAC	CTCGGGTTTC	ATCGTACAGA	CTCTCCTTTC	TTAGACAGCC	TCAAGATCGT
45181	TACACCGGCT	TGCGAAGATG	TACCTTCGCC	TTGAATCEGG	CCCTTGCCAG	CTCGAAGTCCG
45241	ACCACCTGGC	GGGCGGTCTC	CTTCAGGTGC	GACTTCGCCG	ACAGCGGCCG	GACGAACCCG
45301	TAGCTCTTGA	TGTACTCCTC	GAGGTTCGATG	TCGACGTACA	GCGTGACAGG	GACCACCGAC
45361	AAGTCACACC	TCCAATTCGT	GGGGCTTGAT	ETCGTTGGTC	ACGTCTGAGT	CGTTCAGCAG
45421	CGACTGGAAG	TCGGAGTGTG	TCAAGTGTGC	CAAETCATCC	TGCTCGAAGC	GCGCGGGCTC
45481	GTCAATGCCAC	GTCTTCCACT	GGTCGTGGTC	GGCGEGGAAC	CACCTCCGCA	GATCCTTGAT
45541	GGCCTCGTCC	TCGGTGGCGA	AGACGTAGGT	CTCGAGCAGG	TCTCTGTAAT	CGACGGTCAG
45601	CGACCAGACG	GTGATCTTCA	CTCCCCGTTT	ACCTCCGCTT	TGTAGTTTAT	CTCGCGGGTC
45661	TCCTCCTAGT	TGGGTAGCAG	TCGGTTGTAC	TCGTCTGGGC	TGATCTCGCC	AACGATGAAC
45721	TGGCGCATCA	GATTTGCGAC	CGAAGCCGCG	TCCATCCCTT	CGGGAATGGG	CTTGGCGTGG
45781	CCGAAGTGGC	AGTCTCGTGA	CGCCAGCGG	AACCAGAGTT	GGACCTTGTC	CAGTGAGGTC
45841	AGGTGCAGGC	ACTGAAACGT	GATGCCCTCG	AACGGGAAC	CCATCACACC	TCCTGTTTGA
45901	CCTTGACGGT	GTGGCCTGTC	ATTACTTCGT	GGATTGCGAT	GCTGGTGCCG	AACGTCTTTC
45961	GCGTCTCGGC	CTTGAACCTG	GTGGAGCACC	CCGAGCACTT	CGCTTTGAAT	CGCACTAGCA
46021	GTACCAACGC	TTTCTGCAGA	ATCGGCACTT	GCCGCGGTCG	CGGTGTCTGT	TGTCCCGGCG
46081	GGCTTCGCCC	TTCCGTTGAT	CGTCACATGA	CGGAAGCTCG	CCATGCTTGA	TGTGCCATGC
46141	GTCGTGCGCG	ACTTTTCCGC	CGTGCTCGGC	GATGTGCGCT	GCGTCCGGT	ACTCACAGAG
46201	CGGGGAAGCC	GATGCCTCGG	CGATGATCCC	AGGCAGGTTG	CCTAGAACCA	CCGCCAAGCA
46261	CATCAGCAGA	ACGACGTGCC	ACGCCCTCAT	CAGCCCGCCA	GCGCGTGGTT	CATCGCCGCG
46321	TTGCGGCCGT	CGCGCTGACC	GTGGCCTATG	CCGCTGAGGT	CGTACCGGGT	CCGAGGCTTG
46381	ACGTTCTTGG	TGCGAGGATG	CGCCTGGCGC	AGAGCCAGCG	CAGCTCGTTC	CTTGTGCGCT

SUBSTITUTE SHEET

46441	CGGTAGAGCA	CCAACGCTCC	CCCGCCGGCC	GATTCCACGG	CCTTGTTCTC	CTCGGCGGTC
46501	AGGCGTTCCT	TGACGGCCTG	GGCGAAGCCT	GCGATCCACG	ACCGGCGGTA	GCTCTTGAGC
46561	TGGCCAGCGG	TGCTCTTCGG	CTTGTAATCC	CCGGTGTGTG	AGTCGTACTT	GTACCGAGGC
46621	TCGAAAGCCT	GCTCCGGGCG	GACATTCTCA	ACCAGGCGCA	TCATCTGCGG	CTGCATGATC
46681	GACCAGAGGA	ATTGGAGCCT	CTCGATGTGG	CGGGGCACGC	CGTAGACGTA	GATCCGCTGA
46741	CCGCCCGTGA	GGCTGGCGTA	CACCGTCTTG	CAGTGCAGGG	CCTGAGCCAT	GCCGTGCAGC
46801	AACAACGCTT	GTGCGGCAAC	GTACTTGCCG	GTGACGTAGG	TGACCCACTG	GATGGCGTCC
46861	GGCAGGTCGG	TGGTGTCCAA	CCCTTGCTTG	CTCGCCTCGA	CCTGGGCCAT	CTCCAGCCCG
46921	TACTTGGCCA	TCAGCTCGAA	CGCTTTGCGC	TGGAACACAG	CCTCTTCCGG	CGTACCGGCC
46981	ACGTCTTCGG	CCTGGCGCAG	CAGCTTGCGG	ACCTTGCTCT	GCATCTTCTT	CGTCTTGCCG
47041	TCGATCATGG	TCAGTACTCC	TTCTTCCAGT	TGTTCCGGTT	GCCCTTGCCG	GGGCGCTTCA
47101	TCTCTCGCTT	SCGGTTACGG	TGCGGCTGCG	CCGCGTTGGA	GAGACGCAAC	TCGAGCCGTG
47161	CCTTGAGCTG	GTCGCTCATC	TTCTTCACTT	CTTCTGGTTC	AGCGGATCTG	GTCGACGTGG
47221	ATGCAGCCGA	CGCGGTCTGG	CCCGAACTCG	GGAGCGAAGC	CCAGACTTTC	GTCCTCTCTG
47281	CATGGGAACG	CTCGCTGGTC	GAACGTGATT	GGGTGCGCCG	AAGCCTCGTA	TGGATCGGCC
47341	AAGGCCATCG	CTCGGACCGC	TGTAGCGAAT	GCAAGGACGA	CGGTGATCAG	GTGCTTCTTC
47401	ACTCTTCTTC	CCTCCACTTT	TGGTCTGCGA	GAAGCCTTCT	GGCGATCTCG	ATAGGTTCTGA
47461	TCTCAGGAGT	CACTCATCGC	CCTCCAAGAT	CTTCAGGTTG	GCCAGCAGTG	CATTGGCCAC
47521	AGCTCCGATG	TGGCCACCGC	CCTTACTTCC	ACGGCGGGAG	TACTCGCGGT	TCGCGGCGCTG
47581	CATGAAGTGG	AACCTCGGTG	AGCCGTCTCT	GTGAACCCAC	GAGGCTTTCT	CGGCGGGCAG
47641	AGCCCGGTTT	ATCTCCACCG	ACATCGTGAC	GATGATGTGG	TCCCTCTGGA	GCCGAGCCTC
47701	GGTCTCGGCG	TAGTGGGCGG	CTTGGATTAC	TGCGCCTCGT	GTGGTCATGT	CTTCTCTCTC
47761	GGTAGATGTC	AAGCTGTCTG	CACCACTCTT	CGACCGGTAT	CGGTTTGTCA	CAGCCAGCAA
47821	GGATCCCGGC	GTTGCTGCGG	TGATGCCCGT	CCCACAGCGT	CTTTCGGTCC	CTCGAAACCT
47881	CGAGGGGTTT	GAACGGCCAC	TGTTTCGATG	AGTTGAGGAT	GTCCACGACT	TCGTGGACCT
47941	TGGCCAGAA	CTTGCCGGTC	ACGCCCTCCCT	GGTAGTTGTA	GCGGGGCGTG	GTCTGGTAGA
48001	ACTCTTCGAG	CAGTGGTCCG	CTGTCCGCGA	CGGTGCAGTC	GACACGAGCG	CAGGACATGC
48061	AGTCGCTGGC	GCGGAGCTGG	GCAACTTTCAT	CGGTGGTTCAT	GAACGCGGTG	GTACACATCGA
48121	GCCTTTCAGG	TGTATGTCAA	GCGGCGCGGA	CGCCGGAATC	GGAGAGGTAG	ACGCGGTCAG
48181	CTCCACGAA	CGGAGCCTGT	GTGTTGCGGT	GGACGAACTG	GTCGTTCTCG	TAGGGGTTGT
48241	AGGCGATCTT	CGATCCACAG	AAGTCTTGCG	GGAGAAGCGA	GATCAGCTCG	CCTACGATGC
48301	CAGCGTGGAC	CACCTTGCGG	CGCTCGCGCC	GTACCTTGTC	GCGGCGGGCC	GGCCGAACCA
48361	CACCCCTGGC	GTGGGCCAGC	AGGACGTGGC	CGCTGCGGTG	GATGACTCGA	CCCTTGAAGT
48421	CTCCCTCCAA	GGCTTGCAAC	GAGTACACAG	GCCTGCCCTC	GCGGTGCGTG	GCGTGCAGGT
48481	TCTTGTAGAC	GAAGACTCGG	ATCGGCTTGG	GAGTCATGAG	ACCTCCAGTG	TGCGAACGGC
48541	CTTGTAGGCA	CTGATGAGTG	ACGCCCCCGA	CAGCTCGTTA	CCGTGCAGGT	GATACCTGTA
48601	TTTCAGATAC	ACGGCTTGTT	CGACCGGCTT	GTACTCGACC	GAAGTGACCT	CGACAACCAT
48661	CCCGTCGATG	ATCGCGAAGT	CTCCAGCGCG	GAGATGGGTG	GGGAATTTGA	TCTCGGTGTT
48721	GACTACGGTC	ACAGCTTCGA	AACCTCCCAG	GTACCAACGA	ACTTGCCGTT	GCGCTTGATG
48781	TATCCGCTCT	CACCGGGCTC	GTACCAATCG	ACCTCGAACC	CGTAGCGGGC	GGCGCAAGCC
48841	TCGAGGTGGT	CGAGCAGGAC	GCGGCGACCG	GACGCGGTAG	CTTCTCCGGT	CAGCCCCGCTG
48901	TCGTTCTTGC	GGACGATGAG	CTTGAACACT	TGGTGCCTAC	CCTTCTGCGA	TGTCTCGGGA
48961	GATCTCGGCG	AAGACTTTCT	TTGCCACGCG	CACGCCGTCC	CAGGTGATGT	CGAACAGTGC
49021	CTCGTAGAAC	TGGTCTCGCA	AGGCTTCGTT	GCCGTGCGCC	AGCGTTGTGA	CGAGCCGGTC
49081	GATGCGGTCC	TCGTGGAAC	TGTAGACCGA	GTGGTTGTAC	GGCTCAGCCA	TATTGGCGTT
49141	GGTTCGTTTC	ACGTTCTCAA	CCACGATGGC	TTGGAATAGG	TGGTTAACCA	GCTCCTCGGT
49201	CATGTTCTAT	CTCTCCTCAG	TAGTCCGTGT	GCTGGGTCTC	GAAGCCTTCG	AGGTACCCGA
49261	CCTCGTCGTC	GTACGCGCTC	GGGTTGCCGC	GCCAGTCGTC	GCGGAGCCTT	TGACCGCTGG
49321	CGTTGTAGCA	GGCACCACAG	TTCCGGCAGT	CCACATCGCT	CTGGCCGTAG	TAGCGGCAAA
49381	CCTCGCCGCC	GCAGCGTTGG	CAGTCCACAG	CGCTGTAAAC	AGGGATCAGG	AAACCTTGTT
49441	CGTCCGTCTG	ATCAGGGATG	CGTCCGGAAG	TCTTGGCAGG	CATAGTTACT	CCTCATAGAA
49501	ACTCGTGGTT	GATGGCTCGG	TGGGCAGCCT	CGCGGAAGGT	CAGCCCGTCG	TCGTACGCGT
49561	CCCGGTACGT	CCAGTCCGCG	ATGCTCTGGT	AACCAAGACC	AAAGGTCTCT	GTCTGTAGC
49621	CGTCCAGCGC	GGCCATCCAG	GTCTCAAGC	TCATGTCTTC	CCTCACTTCT	TTGTGGTCTGA
49681	GAACAGCACG	TTCCTGCGGC	CGTTGACGCA	CAGACCGCAA	CGGGCACAAG	CCGATCCCTT
49741	GTCGTTGATC	AGGTGATGG	CTTTGTTGTT	CTCCGGGCAG	CGCACCGCCG	TCGGAACCTC
49801	GGCCTTGCTT	TTGGCGAAGC	TGGTGTGAC	GTAGGCGATG	TTGATGCCCT	TGTCTTCCAA
49861	GAAGCGCGCC	ACGTGATGT	TGTGCGGCTG	TGCGCTGAAG	TACAGCGCCA	GGTTGTGAG
49921	CCTCTGCGAG	TGCAGGTAGA	CAGCCGCGGT	CTGAACCTT	GTGTAGGCC	AGAACTGGAC
49981	ATCCGGGTTG	TCGCGGATGA	CTCGACCCCA	AGCGGCCACA	TAGGTGGGGC	TGAAGAAGTC

SUBSTITUTE SHEET

50041 TCCATCCAG TGGATGCGGA ACAGCTTCGG AGCCTTGCGA CGGTCGCAAT CTTTGACGAA
50101 CTCGGCGACC ATCTGGGACA GCAGCGTCAC GGTGTCTGTC AAGTCAGCGT CACGCAACAG
50161 TTCCAGTTG TGCAGCAGGA CCGAGCTGAC AGCCTTGCGA ACTTTCTCCA GCTTGCCGGC
50221 GTAGCACACC TTGGCACAGA AGGCCGTCGC GTCCGGGCAG GAGAAGCCTT GACCGGAGGG
50281 CAGGCCGATG CTGTTGGCGA TACCTACGGT GCGGTTGCCG CCCTTGGTGA CGTGGACGTA
50341 GTTGGTGACC TTGCGGTCGT TCGAACGCTT CAGCTTGGCC ATACCTAGCC TTCCTTCGGT
50401 GGCTGTCAAG TTGTTGGATA CAAAGCGCCC CGAGAGGGAG TCGAACCCCTC ACACCGCGAA
50461 CCGTCGGGGG GCCACCGTGC CTAGTCGATA GAGGTCACTC GACTCTCGTG GACGTAGACC
50521 ACGGTGTTGC CTACGTTCACT CGCGTAGTAC AGGCCATCGG CACCTCGTAG CTTGTGCCGA
50581 ACCGTGCCCG ACGTGGCCGT CATGICTTGG CCCCAGTCGG CGTTAGGTGC CCAGGTGACT
50641 CGCATGGTGA TCCCTTCAGT AGTGGGTGGC TGTCAAGTCA GCGGATACGG ACGTACCCGT
50701 TGCTTCGAGC GACGTAGATC TTGCGTTCGA TGTAAACGCG CTGCTGCTGG TTCATAATCC
50761 TATTCCTTTC GGTGGGTGTC AAGTCTCAGG CCCAGCGACG AGTCGTCCGG CGGGGGCCGG
50821 GCACCTTGGG CCGGTTGGCT CGCGGTCCCT TACGGATGGC GGTGCTACG GTGATCTCTT
50881 CCAACTGGCG TTCAGCCAGG CCGACAGGCC GGGCGTCACC GGGCAGTTCG ATCTTGTAAAT
50941 CGAAGTCAGT CCACCCCTTC AGACCTTCTT CCAGCTCGCG ATCCAAACAGA CGCGGAGCCG
51001 ACAGCTCAGG CGCAACAAAC GGTGTCTTGA CGCTCTCGCG GGCAGTAACC CGAACCTCAC
50161 GGTGCTCAGC GAAGACTGGC ATAGTTCACC CCTTGGTGG ATGTCAAAGCT TGAGCACCAA
51121 AGCTCAGGCG TAGTGGGTAG TCGGGAATCG AACCCGATAG GTTCATAGCC ACGTTCTACG

51181 GCTCAGCCAT AGCTCAGCGA TCATTCCATC GCGCCAAGAG CTACCCTCCC GAATGCCGAA
51241 CCAAAGCTCA GCATTCTGTA GTGTGTATTC TCCCCTGGC TCAGACAGTA TCTATCAGAA
51301 CCTAACACA GGTCTACATT TAGTTATCCG CAGTGCTCGC ACTTTAACGG CATCGAGCTT
51361 CCGCCGACCC TCAGTCTCTT GGCAGCGAAC TAAAGGTTTG AGTCGGGCTG CGGCCCTTCT
51421 CGGTCTTGGC TGATTCTCAC TCTACCGGAT GTTTCGGTGG CTGTCAAGCG GGCCGTTTTG
51481 GTGTTGCAAC GATGCCCTCG TTTAGCGCCG CTGGCGTAAT GCGCTACCCG CCTGATCTCA
51541 CCGGTCCAAG TTGGTGATGC TTGCAGCTTA CCGGATAACC GGGTGGCTGT CAAACCGGAG
51601 AATCTTGCCG CCGGATTTTC ACCGGCACCG GCACGATCCT CTCGGATCCG CCTACCGCCT
51661 TGCTGCTGCG GTCAACAAAG AATGCACTAC TGGCCGGGTG GCTGTCAAGC CCTAATCGCA
51721 AATTGGTGCC CTAGCTGCAG ATATGGCGCG TTCTCGGTGG CTGTAAAGGG CACTACGTGC
51781 CGCTATCCGC TGCTCACGCT GGACAGTCCC GGCAGCCCGT GCCGCGCATA GGCTGCTCAC
51841 TACGTGCCCG GTATCGGCGT TGTGCTGCCG CTGTGCTGGT CGTCGCCCCG TCGCTGTGCG
51901 TGGTCTCGGT GGCATCGCTT GACAGTCGCC CCGCTATCCC CCGTTGCCCG TGGTCAGACG
51961 CTAATCCGCT TATTTGCGAT AGGCTGCTCA CTATCGCATC GGTATGCGTA TCGCTGGTC
52021 ACATATGCGT GTGGTGGTGG TGTGGTGTGC GTGTGTTGC GCTGGTCAGC CGTGTGCGTA
52081 CCGTATCCGC AACTGTGCT TGTGCGTTTG CTGTGTGTCG AGGCCGGCTC TCGCATCGTC
52141 GCATGTGAGC GCGGGTATGG GCGTGTATCG CACGCTTTGC TAGCQCGCTG CCGCGGCGCT
52201 CTCGCATCGC ATCGAGTGTT TGCTGTGCTT CTCATCGTCG CAGGTACAGAA GGGGTAGGGG
52261 GGTTCGCCCT AGGGGTGCGT CCTTGACCGT TCGGTTAT

It is known that during the establishment of lysogeny, the L5 genome becomes integrated into the mycobacterial chromosome and the attachment site (attP). Integration-proficient plasmid vectors have
5 been constructed which efficiently transform both fast-growing and slow-growing mycobacteria through stable integration of the plasmid sequences into the bacterial chromosomal attachment site (attB).

Because the L5 sequence is now known, and
10 because L5 has been previously characterized, the use of transcriptional promoters with this mycobacteriophage may be evaluated efficiently, and host synthesis inhibition may also be evaluated efficiently.

15 Figure 1 represents the genome organization of the entire L5 genome. DNA analysis has indicated that the L5 genome is organized into a right and left arm with the attachment site and integrase at the center of the genome. The integration functions have been
20 successfully employed to construct integration-proficient vectors for mycobacteria.

Part of the L5 genome is not essential for mycobacteriophage growth. It has been demonstrated that all or most of the gene 62-61-60 can be deleted
25 without affecting the cycle of the L5 phage. Therefore, there is a suitable region in the L5 mycobacteriophage for the insertion of reporter

genes. It is critical that reporter genes be inserted into non-essential regions of the mycobacteriophage. Otherwise, the mycobacteriophage will be unable to survive and replicate.

5 The L5 mycobacteriophage may have introduced therein promoter gene 62 fused to reporter gene lacZ, and this reporter mycobacteriophage will be capable of rapid diagnosis of mycobacterial infection and accurate assessment of mycobacterial strain drug
10 susceptibilities.

 Another mycobacteriophage which may be successfully used to produce the reporter mycobacteriophages is the mycobacteriophage TM4. TM4 has been used to construct a first generation reporter
15 mycobacteriophage, and has the ability to discriminate between M. tuberculosis and BCG. A shuttle plasmid may be employed with TM4, and may be useful in the construction of recombinant and other mycobacteriophages. Unlike L5, which is a broad
20 host-range mycobacteriophage, TM4 is a species-specific mycobacteriophage. However, TM4 is not as well characterized as the L5 mycobacteriophage, and therefore it is more difficult to analyze its functions.

25 DS6A is a mycobacteriophage that has been found to be specific for the M. tuberculosis complex of mycobacteria. It has been shown to infect both

M. tuberculosis and BCG. It has been demonstrated that DS6A can infect over 3,000 different types of M. tuberculosis strains. Current efforts are under way to develop DS6A shuttle phasmids containing

5 Firefly luciferase genes as the reporter molecule. It is possible that a combination of different mycobacteriophages may be needed to increase specificity and then increase the ability to distinguish drug susceptibilities. DS6A grows on BCG

10 and M. tuberculosis, but does not grow on M. smegmatis.

In anticipation of the need for a diverse set of mycobacteriophages that can effect a broad or limited range of mycobacterial cells, a total of more than 50 unique mycobacteriophages have been collected

15 and isolated by the inventors. 21 new mycobacteriophages have been isolated from soil samples from India, France, England, Israel, Tunisia, Carville, LA and New York. In addition, another 30 mycobacteriophages from both the Centers for Disease

20 Control in Atlanta and the World Health Organization Phage Reference Laboratory in Amsterdam were collected. The characterization of the nucleic acid content of the phage particles of 30 of these mycobacteriophages have revealed that all of the

25 mycobacteriophages contain double stranded DNA whose genome sizes range from 45 to 100kb as sized on pulsed field gels. Restriction analysis has shown that all

of these mycobacteriophages are different; except that one of the mycobacteriophages from France had a considerable similarity to the L5 mycobacteriophage, which was originally isolated in Japan. The host range of the mycobacteriophages varies greatly, some being able to infect only M. smegmatis and others being able to infect M. smegmatis, BCG and M. tuberculosis, but not M. avium. These mycobacteriophages may be developed into reporter mycobacteriophages and cosmid cloning systems, and may provide a source of useful transcriptional translation initiating sequences, transcriptional terminators, or host-range specificity genes.

In addition, the choice of reporter gene and its method of expression are critical. It is necessary to choose a reporter gene whose product would not normally be found in clinical samples, but whose product is also easily detectable.

Luciferase reporter genes have been used in many diversified biological systems, including E. coli, cyanobacteria, phytopathogenic bacteria and Bacillus. The presence of luciferase reporter genes can be detected by the emission of photons in the presence of a substrate, such as luciferin or decanal. Luciferin and decanal can permeate mycobacteria, and thereby allow for the detection of gene products, such as photons. Since one molecule of

the luciferase gene product can yield 0.85 photons of light, it is the most sensitive biological reporter molecule known. The preferred reporter genes of this invention are luciferase reporter genes, such as the

5 Firefly lux gene (FFlux), the Vibrio fischeri lux genes and the Xenorhabdus luminescens lux genes, as well as the E. coli β -galactosidase (lacZ) genes. Luciferase genes, especially the Firefly lux gene, generate a high amount of luminescence activity. They

10 generate photons, the detection of which is simple and sensitive, using commercially available luminometers that can detect 100-1000 molecules of luciferase with a linear relationship to enzyme concentration. In addition, it is unlikely that clinical samples will

15 contain significant levels of endogenous luciferase activity.

In choosing transcriptional promoters to be introduced into the mycobacteriophages, it is desirable to use strong promoters since this will

20 increase the sensitivity of the system. In addition, it is important that the promoter be active following mycobacteriophage infection. The best promoter candidates currently available are the BCG hsp60 promoter and the L5 gene 62 promoter, which are of

25 comparable strength. The hsp60 promoter gives good levels of luciferase expression from plasmid recombinants, but lower levels of luciferase

expression where the mycobacteriophage is TM4. It is possible that the reason for this is that the hsp60 promoter is shut off by the TM4 enzymes following infection, thus producing only a modest level of luciferase. The gene 62 promoter may behave in a similar manner with the TM4 phage since the gene 62 product is a good candidate for the L5 repressor and is expressed at high levels in the absence of other mycobacteriophage functions. Knowing the sequence of the mycobacteriophage used will help in identifying, characterizing and cloning the appropriate promoter to be used in the reporter mycobacteriophages of this invention.

There are several methods which can be utilized to introduce the reporter genes and transcriptional promoters into mycobacterial species-specific mycobacteriophages. One method is the utilization of shuttle phasmids. When utilizing shuttle phasmid technology, it is necessary to know the sequence of the mycobacteriophage so that the reporter genes are inserted into non-essential regions of the mycobacteriophage. Insertion of reporter genes into non-essential regions permits the mycobacteriophage to survive and replicate. In order to use the shuttle phasmid methodology, it is necessary to first generate a cosmid library of large double-stranded recombinant DNA fragments of

mycobacteriophage. This can be done using cosmid cloning in E. coli. Next, the cosmid library is introduced into the mycobacteria of interest to select for cosmids which have been inserted into

5 non-essential regions of the mycobacteriophage. The shuttle phasmids, which consist of the E. coli cosmid, the reporter genes and mycobacteriophage promoters, may then be characterized. Shuttle phasmids can be propagated in E. coli as plasmids, and propagated in

10 mycobacteria as mycobacteriophages.

A second method of introducing the reporter genes and transcriptional promoters into mycobacteriophages is by homologous recombination or PCR. First, non-essential regions of a

15 mycobacteriophage must be determined. Again, in order to do this, it is necessary to know the sequence of the mycobacteriophage. Consequently, L5 is an ideal phage to use with this method as its genome has already been sequenced and characterized by the

20 inventors. Next, plasmids are constructed wherein reporter genes hooked to transcriptional promoters are flanked by mycobacteriophage non-essential region sequences in mycobacterial plasmids. Then, homologous recombination systems or PCR may be utilized in

25 M. smegmatis or E. coli to perform gene replacement whereby the plasmid constructs containing the reporter genes are put into mycobacteriophages.

A third method of introducing reporter genes and transcriptional promoters into mycobacteriophages is by use of transposons. For example, transposon IS1096 may be utilized. In order to use this methodology, reporter genes and transcriptional promoters are put into transposons, and the transposons containing these reporter genes and transcriptional promoters are delivered on plasmids in mycobacteria. Next, it is necessary to grow up the mycobacteriophages on a strain such as M. smegmatis, which strain contains the transposons. At certain frequencies, the transposons will hop into non-essential regions of the mycobacteriophages, thereby introducing themselves therein. The mycobacteriophages are still viable, and contain the reporter genes and transcriptional promoters.

A fourth method of introducing reporter genes and transcriptional promoters into mycobacteriophages is by debilitated phages packaged into phage heads and tails (phage particles). To utilize this methodology, it is necessary to develop helper phage systems which allow for pieces of DNA containing pac sites to be packaged. These helper phages allow for the synthesis of head and tail genes at will in mycobacteria, prevent themselves from being packaged into phage heads and tails, and facilitate packaging of pacmids into phage heads and tails. Helper phage systems may

be generated from the L5 mycobacteriophage. The genome of the helper phage is put into the mycobacterial chromosome, at which time the mycobacteria are grown up. Next, pacmids which
5 comprise phages which have pac sites, reporter genes, transcriptional promoters and mycobacterial replicons are transformed onto the mycobacterial strain. The production of head and tail proteins may be induced, for example, through an increase in temperature, and
10 the pacmids are then packaged into phage heads and tails. The L5 genome has cohesive (cos) termini. This suggests the possibility of constructing L5 cosmid vectors, which could be packaged through the cos sites into L5 particles either in vivo or in
15 vitro. Then, a large number of genes could be easily and efficiently delivered to mycobacteria.

Packaging into phage heads and tails may also be utilized in a fifth methodology wherein the pacmid is a plasmid. The methodology is similar to the
20 methodology wherein a debilitated phage is used, however, instead of using phage pacmids, the pacmids comprise plasmids which have pac sites, reporter genes, transcriptional promoters, and plasmid replicons.

25 Finally, direct cloning using recombinant DNA techniques in vitro may be used to introduce reporter genes and transcriptional promoters into

mycobacteriophages. This methodology consists of ligating a mycobacteriophage, identifying or introducing unique restriction enzyme sites in non-essential regions of the mycobacteriophage, 5 cleaving the mycobacteriophage with the restriction enzyme sites, and cleaving DNA which encodes the promoter and the reporter gene so that it has the unique sites flanking it on either side. Next, ligation is set up in vitro between the cleaved 10 mycobacteriophage with the unique restriction enzyme sites and the reporter gene cassette. The result is a circular DNA molecule which consists of the mycobacteriophage, the reporter genes and the transcriptional promoters. The circular DNA may then 15 be electroporated directly into mycobacteria.

EXAMPLES

Expression of Reporter Gene lacZ and FFlux in Mycobacteria

A promoter probe vector was constructed which 20 incorporated a truncated E. coli β -galactosidase (lacZ) gene as a reporter probe into a shuttle plasmid vector that replicated in either mycobacteria or E. coli. Random DNA fragments from the three mycobacteriophages L1, TM4 and Bxb1 were cloned into a 25 unique BamHI site immediately upstream of the lacZ gene and screened for their ability to produce β -galactosidase. This established that lacZ could be

used as a reporter gene in the mycobacteria, and identified the DNA sequences which could effectively express foreign genes in both M. smegmatis and M. tuberculosis. β -galactosidase activity could be detected from lysed cells using OMPG, or from unlysed cells using either X-gal or a fluorescent methylumbelliferyl β -galactosidase derivative. The promoter hsp60 gene highly expressed the lacZ gene in both M. smegmatis and BCG.

The FFlux gene was cloned into pMV261 downstream from the hsp60 promoter in plasmid pYUB180 (see Figure 2), which plasmid was shown to express the FFlux gene in M. smegmatis, BCG and M. tuberculosis H37Ra. The expression of the FFlux gene was detected by observing luminescence of mycobacterial clones containing the cloned gene in the dark room, and verified use in photographic film. This demonstrated that the luciferase was expressed in the mycobacteria, and that luciferin, the substrate used, was able to penetrate mycobacterial cell walls and yield photons expressed by the mycobacteria.

Detection of Photons In Mycobacterial Cells Expressing FFlux

The expression of FFlux from the plasmid pYUB180 in M. smegmatis provided a model with which to determine a minimal number of individual cells

detectable with the luciferase assay. M. smegmatis containing pYUB180 were grown in the presence of kanamycin to ensure that every cell contained the plasmid. The cells were diluted 10-fold serially and the amount of luciferase activity was determined using a luminometer. Figure 3 shows that the amount of luciferase activity from 5×10^7 cells approached 10^8 luciferase units, though at this level of activity the luminometer was unable to yield an accurate measurement. However, the activity decreased in a linear manner down to 1200 units for 500 cells. Hence, 5000 cells expressing the FFlux gene can be clearly discerned above the background measurement, which approaches the number of cells that one would expect to observe in clinical samples.

Distinguishing Drug-Resistant Mycobacteria From Drug-Sensitive Mycobacteria Using Luciferase Activity

Since Firefly luciferase activity requires ATP, and ATP is produced only by living cells which are metabolically active, luciferase is a powerful indicator of the metabolic abilities of a bacterial cell. Since anti-tuberculosis drugs are likely to significantly decrease the metabolic activity of a cell, the measurement of luciferase activity should provide a sensitive means of distinguishing drug-resistant mycobacteria from drug-sensitive mycobacteria.

First, the kinetics of the production of luciferase activity of M. smegmatis containing pYUB180 following the addition of streptomycin, isoniazid, ethambutol, rifampicin, ciprofloxacin, novobiocin or cyanide, added at levels that inhibit the growth of M. smegmatis in plate assays, was measured.

As shown in Figure 4, Panel A, the levels of luciferase production were 100 to 1000 times less at eight hours after the addition of the drugs compared to the untreated control.

Next, this approach was used to distinguish drug-resistant from drug-sensitive mycobacteria. The pYUB180 deposit was transformed into streptomycin-resistant or novobiocin-resistant M. smegmatis mutants. Photon production by the drug-sensitive parent was compared to the streptomycin-resistant or novobiocin-resistant mutants. The drug-resistant mutants continued to produce luciferase activity levels comparable to the untreated parent in the presence of the appropriate antibiotic. In addition, the drug-resistant mutants produced 100 to 1000 times more luciferase activity than the drug-sensitive parent (see Figure 4, Panels B and C). Hence, a luciferase-based assay may be used to determine mycobacterial drug susceptibility.

Construction of TM4 Reporter Mycobacteriophages and
Detection of Photons Following TM4::lux Infection

The first vectors developed to introduce recombinant DNA into mycobacteria were shuttle phasmid
5 phage vectors. Shuttle phasmids have the ability to replicate in E. coli as cosmids and then replicate in mycobacteria as phages. Shuttle phasmids of TM4 which contained the FFlux and lacZ genes transcribed from hsp60 and L1 promoters, respectively, were constructed
10 (see Figure 5).

A deposit of the shuttle phasmid (reporter mycobacteriophage) phAE39 which contains mycobacteriophage TM4, cosmid pYUB216, reporter gene FFlux and promoter hsp60, was made with the American
15 Type Culture Collection on January 12, 1992 and catalogued as ATCC #75183. When the TM4::lux shuttle phasmid phAE39 was mixed with M. smegmatis cells, luciferase activity could be detected within 15 minutes of incubation, and continued to increase
20 slightly over the next 4 hours (see Figure 6). These results show that the TM4::lux mycobacteriophage is capable of introducing the FFlux gene into mycobacterial cells, and that the FFlux gene can be expressed in mycobacteriophage-infected cells. Figure
25 7 represents a flow chart for cloning different promoters into the TM4::lux shuttle phasmid phAE39.

A deposit of the shuttle phasmid (reporter

mycobacteriophage) phAE37 which contains mycobacteriophage TM4, cosmid pYUB216, reporter gene lacZ and promoter L1, was made with the American Type Culture Collection on _____, 1992 and catalogued as ATCC #_____. The TM4::lacZ mycobacteriophage formed bright blue plaques when plated on media containing X-gal.

Construction of the L5 Reporter Mycobacteriophage

Strategies for construction of the recombinant L5 mycobacteriophage may be investigated. The possibility of using the shuttle phasmid approach starting with L5 deletion derivatives, in which the size of the genome has been reduced, may also be explored. Initially, the largest gene 62 deletion available should be used. However, other deletion derivatives in which more of the gene 62-61-60 segment is lost should also be isolated. Another approach would be to attempt to introduce genes by homologous recombination with plasmids. Still another approach would be to transpose lux genes onto L5 using either the mini-Mu in vitro transposition system or a mycobacterial transposon such as IS1096.

Recombining reporter genes from recombinant plasmids onto L5 using a double recombination event may also be performed. This involves first constructing a recombinant plasmid that carries a reporter gene (lacZ may be more suitable) inserted

into gene 62 such that both the upstream and downstream parts of gene 62 are present. Advantages of this approach are that lacZ can be easily detected in agar media, that gene 62 is not an essential gene, and that lacZ is efficiently expressed from a promoter immediately upstream of gene 62. An L5 mycobacteriophage lysate may be prepared by growth of the plasmid-containing strain and recombinant mycobacteriophage progeny identified by plating the lysate on wild-type M. smegmatis for individual plaques on agar containing the indicator X-gal.

This recombination approach may be expanded to introduce other gene or DNA segments of the L5 genome. For example, it should be possible to add luciferase genes from FFlux in an identical manner, provided that packaging limits are not exceeded. In addition, inclusion of polylinker containing restriction enzyme sites unique for L5 would open the way for construction of L5 recombinants in vitro. Similar genetic strategies may be used to systematically reduce the size of the L5 genome by deletion of non-essential sequences.

Transposition offers an alternative method for the construction of reporter mycobacteriophages. A transposition system which is available is the mini-Mu in vitro transposition system. This is a defined biochemical reaction in which a mini-Mu transposon

carrying the desired gene is transposed onto the phage genome using purified MuA and MuB proteins. Similar transposition experiments have been tried with L5, but few L5 mini-Mu derivatives have been isolated. It is possible that this is due to the relatively large size of the transposon used. It is necessary to first construct a small Mu transposon which contains the reporter gene, a promoter and the two Mu in order for these experiments to be successful.

10 Development of L5 in vivo
 and in vitro Packaging Systems

g cosmids and packaging systems provide the efficiency of mycobacteriophage infection with the ability to inject large segments of non-mycobacteriophage DNA. Analogous mycobacterial systems would overcome packaging constraints encountered with recombinant mycobacteriophage genomes and allow the introduction of multiple copies or types of reporter genes into mycobacteria, potentially enhancing the sensitivity of the assay. In addition, they would help overcome any problems with host synthesis inhibition.

20 The development of L5 cosmids and packaging systems is dependent on the finding that the L5 genome contains cohesive termini. The g paradigm suggests that a relatively small region of DNA (approximately 500bp) around the cos site (in the ligated form) is

necessary to promote packaging. The first series of experiments with L5 would therefore be to identify the segment of the genome required for packaging by constructing a series of plasmids containing the L5 cos site and surrounding sequences. Cos activity may be determined by preparation of an L5 lysate on plasmid-containing M. smegmatis strains, followed by the identification of antibiotic-resistant transductants in the lysate, by transduction of M. smegmatis. This assay assumes that plasmid multimers of a total size of approximately 50kb are present in the cell and will be packaged. Although the presence of such multimers has not been demonstrated directly, they are likely to be generated by the homologous recombination system of M. smegmatis. If this assay should fail, cosmid vectors which contain both L5 g cos sites may be constructed. Insertion of 40-45kb of DNA (as in the construction of cosmid libraries) followed by g packaging in vitro and infection with E. coli will generate 50kb sized molecules containing L5 cos site. These should be isolated from E. coli and introduced by electroporation into M. smegmatis. Assuming that one of these approaches is successful, it would then be possible to define a small segment of L5 DNA required for packaging.

The construction of in vivo cosmid packaging

systems is a particularly attractive idea since it has proven very useful in E. coli. Thermoinducible lysogens of L5 may be suitable for in vivo packaging of L5 cosmids without further modification, since
5 prophage excision may be a temperature-sensitive event. Efficient packaging of extrachromosomal cosmids present in the lysogen may be achieved by simple induction and growth at 42°C.

It is possible that some process other than
10 excision is temperature-sensitive in lysogen induction. If so, it will be necessary to further debilitate the prophage in order to prevent DNA packaging of the prophage. There are a variety of ways to accomplish this. For example, the excise gene
15 itself could be deleted (using a recombination strategy similar to that described above) such as to prevent excision. Another approach is to damage the cohesive termini (by exonucleolytic digestion) of an L5 thermoinducible derivative and construct a
20 defective lysogen. A combination of approaches may be desirable, since even if prophage excision is a temperature-sensitive process, the destruction of cos might effectively reduce the background of spontaneous mycobacteriophage release.

25 Construction of in vitro packaging systems will follow similar lines. Extracts may be prepared from thermoinducible strains with non-packagable

prophages and assessed for their ability to package exogenously added L5 cosmid or mycobacteriophage DNA. Optimization of conditions should follow both empirical biochemical approaches and the well-established ϕ systems. For example, it may be necessary to supplement the extracts with purified mycobacteriophage products such as the terminase or the tape-measure analogues (genes A/Nu and H of ϕ respectively), neither of which have yet been identified.

Construction of Novel Shuttle Phasmids From Any Mycobacteriophage

Although mycobacteriophages L5 and TM4 can be used in the development of diagnostic luciferase and β -galactosidase shuttle phasmids, there may be other mycobacteriophages, such as the mycobacteriophage DS6A which only infects BCG and M. tuberculosis strains, that might prove to have a more useful host range for clinical isolates. Diagnostic luciferase mycobacteriophages from these other mycobacteriophages may be developed by using the shuttle phasmid methodology described herein that has been proven successful for constructing mycobacteriophage vectors from both TM4 and phage L1.

Isolate Mycobacteriophage L5 and TM4 Mutants to Infect the Maximum Number of Clinical Isolates

For the diagnostic luciferase mycobacteriophage system to have maximal use in the clinical laboratory,

it will be essential that to develop a set of diagnostic mycobacteriophages that can efficiently infect any clinical isolate and possibly distinguish M. tuberculosis from M. avium and BCG. Both mycobacteriophages TM4 and L5 appear to have the ability to infect a large number of M. tuberculosis isolates. TM4 is very closely related to phage 33D, a mycobacteriophage that has been found not to infect every M. tuberculosis isolate used to define the mycobacteriophage typing schemes for M. tuberculosis isolates. However, this mycobacteriophage does not infect BCG. TM4 has been found to be almost identical by DNA hybridization and restriction analysis to 33D, and it shares the host-specificity with 33D in that it infects M. tuberculosis, but fails to infect BCG. mycobacteriophage L5 appears to share the same receptor as mycobacteriophage D29 which receptor has been previously shown to infect a very large number of M. tuberculosis isolates. L5, unlike 33D or TM4, infects all three morphotypes of M. avium including a wide range of serovariants.

If L5 or TM4 are found not to infect certain M. tuberculosis isolates, it may be possible to isolate mutants of these mycobacteriophages which plaque on the particular isolate. The inability to plaque on a particular isolate could result from the lack of a mycobacteriophage receptor or be the result

of lysogenization of the isolate with a homoimmune phage. Phage mutants with altered host range specificities or mutants which no longer bind a repressor (equivalent to virulent mutant of g) have been isolated in other systems. Variants of TM4 which can efficiently infect BCG have been isolated at frequencies of 10^7 . Previous work has demonstrated that 33D, similarly to TM4, can not adsorb to BCG cells. Host-range variants of TM4 which not only plaque BCG, but also still plaque M. tuberculosis have been isolated. Similar strategies for M. tuberculosis isolates which are uninfected by L5 or TM4 may be used.

Detecting the Presence of M. tuberculosis in Clinical Samples

The combined sensitivities of luciferase and mycobacteriophage infections should permit the detection of previously undetectable levels of M. tuberculosis cells in sputum, blood samples, or cerebral spinal fluid. A number of preliminary studies to optimize the detection of M. tuberculosis cells in a variety of body samples will be performed.

Detecting M. tuberculosis Grown In Primary Human Macrophages and Macrophage Cell Lines

As a model system for optimizing detection of M. tuberculosis in infected monocytes and macrophages, primary human monocytes which have been purified by adherence for 1 hour or primary macrophages which have

been cultured for 6 days in microwells will be infected with M. tuberculosis H37Ra at varying multiplicities. The number of cells initially infected will be determined microscopically, and then
5 at various periods of time from 2 hours to 30 days, the cells will be lysed by non-ionic detergent NP40 which has no effect on viability of mycobacteria, concentrated by centrifugation, plated for viable organisms and infected with the luciferase plasmids.
10 Quantitative studies at different moi's and with varying numbers of infected cells will indicate how few bacilli/cell and bacilli/specimen can be detected.

The inability of M. tuberculosis cells isolated from macrophages to be infected with diagnostic
15 shuttle phasmids could result from either the absence of the expression of the mycobacteriophage-receptor or the masking of the receptor with a membrane from a phagosome of the macrophage. The level of expression of phage receptors may be regulated by the environment
20 in which the host cell is grown. For example, the λ repressor of E. coli is induced by maltose and repressed by glucose. Studies to identify the receptors for mycobacteriophage L5 have been initiated. Similar studies for mycobacteriophage TM4
25 will also be performed. By identifying the genes encoding the receptor, it is possible to assay gene repression of the mycobacteriophage receptor of

M. tuberculosis cells when grown in macrophages by hybridization for the mRNA synthesis. If the receptor is not expressed in macrophages, it may be necessary to use a mycobacteriophage which recognizes a receptor that is constitutively expressed.

If the receptor is masked by a membrane of the macrophage, the cells isolated from macrophages may be treated with a variety of different detergents to find a treatment that would allow infection of the M. tuberculosis cells with the mycobacteriophages. Again, it may be necessary to cultivate the detergent-treated macrophages in broth for a few generations to gain expression of the receptors. The assays to determine the infectability of macrophages from mycobacteria include not only the luciferase assay for the TM4::lux mycobacteriophages, but also infectious centers assays in which free mycobacteriophages are removed and mycobacteriophage-producing cells are scored by a mixed plating on a lawn of M. smegmatis. This assay would be useful since infectability can be scored even if there are insufficient M. tuberculosis cells to form a bacterial lawn. It is important to re-evaluate the host range specificities of all of the mycobacteriophages in this assay. Free mycobacteriophages can simply be removed through the use of specific anti-mycobacteriophage antibodies.

Detecting M. tuberculosis in Sputum Samples

Sputum from a patient infected with M. tuberculosis contains a mixture of mucopolysaccharide, free M. tuberculosis cells, macrophages containing M. tuberculosis cells and a variety of cellular debris. Sputum samples from patients thought to have pulmonary tuberculosis may be used for a study in which various numbers of M. tuberculosis cells are added to sputum samples found to have no or few organisms by acid-fast staining. A variety of methods can be used to treat sputum samples so as to liquify the mucous and decontaminate the specimen under conditions in which all bacteria other than mycobacteria are killed. Because of the specificity of the phasmids, decontamination may not be as important as preserving the mycobacteriophage receptors. Nonetheless, the sputum samples may be treated initially with 2% w/v NaOH for 30 minutes at 37°C or with 0.5% N-acetyl cysteine + 1% NaOH. Alternatively, the sample may be treated with a variety of hydrolytic enzymes, such as collagenase, to help dissolve the sputum sample. If mycobacteriophage receptors are carbohydrates possibly sensitive to these conditions, other conditions may be utilized or the cells will be cultured 3-16 hours to allow recovery of infectivity before mycobacteriophage infection.

Detecting Mycobacteria In Blood Samples

Tuberculosis has been known to have a bacteremia. If the sensitivity necessary to detect 100 to 200 M. tuberculosis cells in a ml of sample can be obtained, levels of bacteremia in tuberculosis patients which were not previously observable may be observed. White cells should be purified over Ficoil-hypaque and lysed with 2% NP40, 1% SDS or freeze-thawing in the presence of DNase to liberate intracellular mycobacteria. The pellet should then be infected with the diagnostic luciferase mycobacteriophage, or if only few organisms are present they can be concentrated by filtration onto filters, and filter areas cut out and infected.

15 Assuring Specificity On a Variety of Clinical Isolates and Species; Assessment of False Positives and Negatives

The luciferase assay may be optimized such that positive correlations of M. tuberculosis infections as indicated in the clinical lab may be obtained. The recombinant mycobacteriophages may be tested to ascertain the range of specificity that they have for other mycobacteria, and for the closely related genera Norcardia, Corynebacterium, and Actinomycetes strains. These strains may be obtained from the ATCC. A number of blinded tests including negative controls, M. tuberculosis-infected patients, samples

from patients infected with M. avium, and samples infected with other non-mycobacterial pathogens may be performed to ascertain the range of specificity.

5 The ability to rapidly assess the
susceptibilities of M. tuberculosis isolates to
isoniazid, ethambutol, rifampicin, pyrazinamide and
other antibiotics will have a major impact on the
treatment of tuberculosis patients. After the
isolation of M. tuberculosis cells from a sputum
10 sample, which may take several weeks, the assessment
of drug-susceptibilities may take an additional 2 to 9
weeks. Diagnostic reporter mycobacteriophages may
allow for evaluations of drug-susceptibilities at the
time a sputum sample is collected. Alternatively,
15 this approach would shorten the time necessary to
assess drug-susceptibilities of purified
M. tuberculosis colonies grown up from clinical
samples.

Luciferase Assays for M. tuberculosis
Cells in the Presence of Drugs

20 The results of the experiments suggest that by
using luciferase as an indicator for the metabolic
ability of the cell, it may be possible to define
conditions which will enable us to distinguish
drug-resistant mycobacteria from drug-sensitive
25 mycobacteria. To test this hypothesis, isolated

mutants of M. tuberculosis H37Ra which are resistant to isoniazid, rifampicin, ethambutol, or pyrazinamide would be used to generate a set of cogenic mutants. These independent mutants and the parent strains would be transformed with pYUB180. Luciferase activity will be assessed in the presence and absence of drugs in order to determine the optimal conditions for distinguishing between drug-resistant and drug-sensitive cells. It is quite possible that the window of time to observe differences for different drugs could vary and require different incubation times for each drug.

The choice of the promoter for expressing luciferase may provide a needed parameter to more readily assess drug action. For example, in the case of E. coli, gyrase promoters are greatly stimulated in the presence of gyrase inhibitors.

Clinical isolates of M. tuberculosis may be transformed with PYUB180 and tested for luciferase activity in the presence and absence of drugs. The luciferase assays with mycobacteriophage infections with lux mycobacteriophages on in vitro-grown M. tuberculosis cells will first be optimized, and then extended to M. tuberculosis cells grown in macrophages or isolated from sputum samples.

Critical Assessment of Drug-Susceptibility Testing

As for the detection of M. tuberculosis from clinical samples, the luciferase assay may be optimized so that the drug-susceptibility patterns for any clinical isolate may be obtained. It may be possible to add diagnostic mycobacteriophages to a single clinical specimen, aliquot the mixture into various tubes and add antibiotic drugs. Thus every experiment would have an internal control and each drug-treated sample could be compared to an untreated control. The critical parameter to conclude drug-resistance or sensitivity lies in the comparison.

Although the invention herein has been described with reference to particular embodiments, it is to be understood that these embodiments are merely illustrative of various aspects of the invention. Thus, it is to be understood that numerous modifications may be made in the illustrative embodiments and other arrangements may be devised without departing from the spirit and scope of the invention.

WHAT IS CLAIMED IS:

1. A method of producing mycobacterial species-specific reporter mycobacteriophages which comprises introducing reporter genes and transcriptional promoters into the genomes of mycobacterial species-specific mycobacteriophages wherein upon incubation with the mycobacteria for which said reporter mycobacteriophage is specific, the reporter genes of said reporter mycobacteriophage will express a gene product which is detectable.
2. The method according to Claim 1 wherein the reporter genes and transcriptional promoters are introduced into the mycobacteriophages by shuttle phasmid technology.
3. The method according to Claim 1 wherein the reporter genes and transcriptional promoters are introduced into the mycobacteriophages by homologous recombination or PCR.
4. The method according to Claim 1 wherein the reporter genes and transcriptional promoters are introduced into the mycobacteriophages by transposon technology.
5. The method according to Claim 1 wherein the reporter genes and transcriptional promoters are

introduced into the mycobacteriophages by debilitated phages packaged into phage heads and tails.

6. The method according to Claim 1 wherein the reporter genes and transcriptional promoters are introduced into the mycobacteriophages by plasmids packaged into phage heads and tails.

7. The method according to Claim 1 wherein the reporter genes and transcriptional promoters are introduced into the mycobacteriophages by recombinant DNA techniques.

8. The method according to Claim 1 wherein the mycobacteria is M. tuberculosis.

9. The method according to Claim 1 wherein the mycobacterial species-specific mycobacteriophage is L5, TM4 or DS6A.

10. The method according to Claim 1 wherein the reporter genes are luciferase genes or the β -galactosidase gene.

11. The method according to Claim 10 wherein the luciferase genes are selected from the group consisting of Firefly lux gene, Vibrio fischeri lux genes, Xenorhabdus luminescens lux genes and lacZ genes.

12. The method according to Claim 1 wherein the transcriptional promoter is hsp60 or the L5 gene 62 promoter.

13. The method according to Claim 1 wherein the gene product is photons.

14. The method according to Claim 1 wherein the gene product is made detectable by contacting said
5 gene product with a substrate.

15. The method according to Claim 14 wherein the substrate is luciferin or decanal.

16. The mycobacterial species-specific reporter mycobacteriophage produced by the method of
10 Claim 1.

17. A mycobacterial species-specific reporter mycobacteriophage comprising a mycobacterial species-specific mycobacteriophage which contains in its genome reporter genes and a transcriptional
15 promoter, wherein the reporter genes express a gene product upon incubation with the mycobacteria for which the reporter mycobacteriophage is specific.

18. The mycobacterial species-specific reporter mycobacteriophage according to Claim 17
20 wherein the mycobacteria is M. tuberculosis.

19. The mycobacterial species-specific reporter mycobacteriophage according to Claim 17 wherein the mycobacterial species-specific mycobacteriophage is L5, TM4 or DS6A.

20. The mycobacterial species-specific reporter mycobacteriophage according to Claim 17

wherein the reporter genes are luciferase genes or the β -galactosidase gene.

21. The mycobacterial species-specific reporter mycobacteriophage according to Claim 20
5 wherein the luciferase genes are selected from the group consisting of Firefly lux gene, Vibrio fischeri lux genes, Xenorhabdus luminescens lux genes and lacZ genes.

22. The mycobacterial species-specific
10 reporter mycobacteriophage according to Claim 17 wherein the transcriptional promoter is hsp60 or the L5 gene 62 promoter.

23. The mycobacterial species-specific reporter mycobacteriophage according to Claim 17
15 wherein the gene product is photons.

24. The mycobacterial species-specific reporter mycobacteriophage according to Claim 17 wherein the gene product is made detectable by contacting said gene product with a substrate.

20 25. The mycobacterial species-specific reporter mycobacteriophage according to Claim 24 wherein the substrate is luciferin or decanal.

26. A method of diagnosing a mycobacterial disease which comprises incubating a sample which may
25 contain myco- bacteria with mycobacterial species-specific mycobacteriophages which contain

reporter genes and transcriptional promoters in their genomes, wherein the reporter genes produce a gene product upon incubation with the mycobacteria for which the mycobacteriophage is specific, and wherein
5 the gene product is detectable.

27. The method according to Claim 26 wherein the mycobacterial disease is tuberculosis.

28. The method according to Claim 26 wherein the mycobacteria is M. tuberculosis.

10 29. The method according to Claim 26 wherein the mycobacterial species-specific mycobacteriophage is L5, TM4 or DS6A.

30. The method according to Claim 26 wherein the reporter genes are luciferase genes or the
15 B-galactosidase gene.

31. The method according to Claim 30 wherein the luciferase genes are selected from the group consisting of Firefly lux gene, Vibrio fischeri lux genes, Xenorhabdus luminescens lux genes and lacZ
20 genes.

32. The method according to Claim 26 wherein the transcriptional promoter is hsp60 or the L5 gene 62 promoter.

33. The method according to Claim 26 wherein
25 the gene product is photons.

34. The method according to Claim 26 wherein

the gene product is made detectable by contacting said gene product with a substrate.

35. The method according to Claim 34 wherein the substrate is luciferin or decanal.

5 36. The method according to Claim 26 wherein the sample is blood or sputum.

37. A method of assessing drug resistance of a mycobacterial strain which comprises:

10 (a) incubating a sample which contains a myco- bacterial strain with mycobacterial species-specific mycobacteriophages which contain in their genomes transcriptional promoters and reporter genes which produce gene products;

15 (b) adding an anti-mycobacterial drug to the incubation; and

(c) detecting whether the gene product is present in the sample, such presence indicating drug resistance of the mycobacterial strain.

20

38. The method according to Claim 37 wherein the mycobacterial strain is a strain of M. tuberculosis.

39. The method according to Claim 37 wherein
25 the mycobacterial species-specific mycobacteriophage is L5, or TM4 or DS6A.

40. The method according to Claim 37 wherein the reporter genes are luciferase genes or the β -galactosidase.

5 41. The method according to Claim 40 wherein the luciferase genes are selected from the group consisting of Firefly lux, gene, Vibrio fischeri lux genes, Xenorhabdus luminescens lux genes and lacZ genes.

10 42. The method according to Claim 37 wherein the gene product is photons.

43. The method according to Claim 37 wherein the transcriptional promoter is hsp60 or the L5 gene 62 promoter.

15 44. The method according to Claim 37 wherein the anti-mycobacterial drug is selected from the group consisting of streptomycin, isoniazid, ethambutol, rifampicin, ciprofloxacin, novobiocin and cyanide.

20 45. The method according to Claim 37 wherein the gene product is made detectable by contacting said gene product with a substrate.

46. The method according to Claim 45 wherein the substrate is luciferin or decanal.

47. The method according to Claim 37 wherein the sample is blood or sputum.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/00913

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) : C12N 07/00; C12P 21/06; C12Q 01/66

US CL : 435/ 235.1, 69.8, 8

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/ 235.1, 69.8, 8

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
medline, dialog, aps

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 4,861,709 (Ulitzer et al.) 29 August 1989, see column 8, 14-17, and 28.	1-47
Y	Revista Cubana de Medicina Tropical, Volume 41, No. 2, issued 1989, M.C.A. Jimenez et al., "Phage Typing Marker Study of Mycobacterium-Tuberculosis Strains from Ethiopia Preliminary Report", pages 192-199, see abstract.	1-47
Y	Nature, Volume 351, issued 06 June 1991, C.K. Stover et al., "New Use of BCG for Recombinant Vaccines", pages 456-460, see entire document.	1-47

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	T	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be part of particular relevance	X	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	Y	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	&	document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means		
"P" document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

15 APRIL 1993

Date of mailing of the international search report

21 APR 1993

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Authorized officer

KATHLEEN L. CHOI

Facsimile No. NOT APPLICABLE

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/00913

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Folia Microbiology, Volume 36, No. 5, issued 1991, J. Konicek et al., "Gene Manipulation in Mycobacteria", pages 411-422, see pages 415-417.	1-47
Y	Nature, Volume 327, issued 11 June 1987, W.R. Jacobs Jr. et al., "Introduction to Foreign DNA into Mycobacteria using a Shuttle Plasmid", pages 532-535, see entire document.	1-47
Y	Proceedings of the National Academy of Sciences, Volume 88, issued April 1991, M.H. Lee et al., "Site-specific Integration of Mycobacteriophage L5: Integration-proficient Vectors for Mycobacterium smegmatis, Mycobacterium tuberculosis, and bacille Calmette-Guerin", pages 3111-3115, see entire document.	1-47
Y	Journal of General Virology, Volume 26, No. 1, issued January 1975, J.A. Hewitt, "Miniphage - a Class of Satellite Phage to M13", pages 87-94, see abstract.	5, 6
Y	Journal of Bacteriology, Volume 149, No. 3, issued March 1982, M.J. Orbach et al., "Transfer of Chimeric Plasmids among Salmonella typhimurium Strains by P22 Transduction", pages 985-994, see entire document.	5, 6
Y	Zentralbl. Veterinaermed., Reihe B, Volume 25, No. 5, issued 1982, R. Weiss et al., "Resistance Testing of Bacteria by Firefly Bioluminescence. A Rapid Test", pages 359-71, see abstract.	37-47
Y	Fortschr. Veterinaermed., Volume 35, issued 1982, R. Weiss et al., "Bioluminescent Methods to Test the Antibiotic Sensitivity of Bacteria", pages 323-328, see abstract.	37-47

